Etude du régime alimentaire des carnivores par des techniques moléculaires

Wasim Shehzad

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biodiversité, Ecologie et Environnement**

Arrêté ministériel : 7 août 2006

Présentée par

Wasim SHEHZAD

Thèse dirigée par **Pierre TABERLET** et
codirigée par **François POMPANON**

préparée au sein du **Laboratoire d'Ecologie Alpine (LECA)**
dans **l'École Doctorale Chimie et Science du Vivant**

DNA-based diet analyses in carnivores

Thèse soutenue publiquement le **14 décembre 2011**,
devant le jury composé de :

Prof. William O.C. SYMONDSON
Prof. Cardiff University, UK (Rapporteur)
Dr. Luca FUMAGALLI
MdER, Université de Lausanne, Suisse (Rapporteur)
Dr. Irène TILL-BOTTRAUD
DR CNRS, France (Présidente, Examinatrice)
Dr. Pierre TABERLET
DR CNRS, France (Directeur de thèse)
Dr. François POMPANON
MCF HDR, UJF, France (Co-directeur de thèse)

To my family and my all teachers

Table of contents

List of Abbreviations

ANP	Ayubia National Park
CGNP	Chitral Gol National Park
eDNA	environmental DNA
DNA	Deoxyribose Nucleic Acid
DPO	Dual Priming Oligonucleotide
mt-DNA	mitochondrial DNA
LNA	Locked Nucleic Acid
NGS	Next Generation Sequencing
NTS	Nouvelles Techniques de Séqiençage
PCR	Polymerase Chain Reaction
SLT	Snow Leopard Trust

List of Figures

Preface

This thesis manuscript presents the results of my research work conducted over a three years period (2008-2011) during my stay at the Laboratoire d'Ecologie Alpine (LECA), Université Joseph Fourier, Grenoble. This work complements the approach carried out by Alice VALENTINI on herbivores diets in the same laboratory. The main aim of my study was to develop a DNA-based universal method to study carnivores diet using environmental samples, and to apply this approach to several carnivore species. The main focus was on the diet of endangered and elusive wild felid species of Asia. Studying carnivores diet by characterizing the prey DNA present in its feces was a real challenge when both prey and predator belong to a close taxonomic group. In this case, the universal primers used, will amplify both predator and prey DNA, with a bias in favor of predator's DNA sometimes until the absence of prey DNA in the PCR product. Thus, the main obstacle in studying carnivores diet in such cases was to control the amplification of predator DNA to a limited extent and using universal primers amplifying all potential prey items present in non-invasive sample. My early discussions with Dr. Pierre TABERLET dealt with several approaches such as the use of restriction enzymes targeting predator's DNA. But, one thing which impeded me to work on this approach was that the restriction enzymes which cleave the predator DNA may also cut some prey sequences belonging to same or closer taxonomic group. A second obstacle dealt with the finding of a restriction site, which is not always evident. However, the use of oligonucleotide homologous to a part of the predator DNA and blocking its amplification appeared to be a solution. We started with an artificial mixture combining bear and frog tissue DNA in different proportions. Then we shifted towards the analysis of feces samples from target carnivore species. This work involved the use of Next Generation Sequencing (NGS) technologies. I started with the comparison of bands, obtained on agarose gel due to different proportions of predator to prey DNA. Then, it took three 454 (Roche) sequencing runs, in order to validate the effectiveness of blocking oligonucleotides in various experimental conditions. Finally a large NGS run using the Illumina-Solexa technology was designed including the sequencing of a large number of samples.

This manuscript consists of six chapters. Chapter one is a summary if the whole thesis is in French. Chapter two is an introductive presentation of the context of the

study. It includes a brief review about DNA barcoding and NGS, their potentials and implications in various ecological studies including diet analyses. This chapter mainly focuses on the importance of studying the diet of elusive and endangered carnivore species and I have discussed the perspectives and limitations of various methods from conventional to PCR based techniques proposed till to date.

Chapter three mainly deals with the methodological aspect of my PhD study. It presents the blocking oligonucleotide approach primarily applied for assessing the diet of the leopard cat from its feces by amplifying prey DNA using universal primers for vertebrates.

Chapter four highlights the relevance of diet assessment in the context of human-wildlife conflict, by studying the common leopard population of Ayubia National Park Pakistan. A conflict arises when human population living in the proximity of wildlife relies upon the same resources. Our study confirmed that the common leopard population studied mainly depends upon domestic livestock for its survival, and we have make different suggestions to effectively mitigate these conflicts.

In chapter five we apply our approach to a highly endangered and cryptic carnivore, the snow leopard, a case where the study of feces is required due to its remote habitat and cryptic nature. One of the major threats to snow leopard population throughtout its range, is the retribution killing due to livestock depredation. This study aims to verify the proportion of livestock presence, moreover, the share of nearly threatened argali sheep in the diet of snow leopard population from Tost Mounatins, South Gobi, Mongolia. We illustrated that this predator mainly depends upon wild ungulates, which seem to represent affluent population in Mongolia. These results will allow resolving various conservation issues related with this cat.

Chapter six concludes on the overall accuracy of the approach developed for assessing the diet of carnivores, and on its outcomes in the conservation of endangered species studied. I have also highlighted potential limitations of this technique and possible solutions to overcome them. Finally, I give some recommendations for the implications of this approach to other carnivores diet.

Acknowledgments

First of all, I owe my deepest and sincere gratitude to my PhD supervisor, **Pierre TABERLET**, for choosing me to work on carnivores diet analysis project. He is the man entered me into the new horizon of molecular ecology and has been a continuous source of inspiration and encouragement throughout my PhD. His foresight and wisdom about the science is really matchless. I am also especially indebt to **François POMPANON**, my co-supervisor for his ever-helping support, guidance, and his availability throughout the course of this study despite of his busy schedule. Especially, for his valuable suggestions and guidelines on this Manuscript. His help in translating the French chapter of this Manuscript.

I feel myself incredibly lucky to be the part of the research group of **LECA**. I learnt a good deal from here. I am thankful to direction of LECA especially **Pierre TABERLET** and **Irène TILL-BOTTRAUD** for maintaining very nice working environment in the laboratory, without any hierarchy. I am especially thankful to **Alice VALENTINI,** for her ever support and answers to even of my stupid questions. **Christain MIQUEL** and **Carole POILLOT** for make me aware of 'Power of pipetting', their help during the experimental phases and being a viable source of French language. I found Christian MIQUEL at your help whatever and whenever one needs. Carole POILLOT for her help in translating some part of this Manuscript in French. **Delphine RIOUX** and **Ludovic GIELLY** for their help and support. **Eric COISSAC** for introducing his magic tools of bio-informatics. **Aurélie BONIN,** for her valuable comments and suggestions for the thesis manuscript.

I am extremely thankful to all my field collaborators, arranging field samples for this study, including **Ali NAWAZ,** Assistant Professor Quaid-e-Azam University, Islamabad, for arranging common leopard and leopard cat samples from Pakistan. **Tom McCARTHY** from Snow Leopard Trust/Panthera, for making possible of fecal samples of snow leopard from Mongolia. **Eric MARBOUTIN, Christophe DUCHAMP** and **Pierre-Emmanuel BRIAUDET** from **ONCFS** for arranging fecal samples of wolf from French Alps.

I owe my thanks to administrative secretaries of LECA, **Florence SAGNIMORTE** and **Kim PLA** for always making the administrative stuff easier.

Olivier LONTIN, for his help despite his quite precisions. My friends of Bureau 308**,** especially **Carole POILLOT, Guillaume TETREAU, Frédéric LAPORTE, Marti BOLEDA, Bruno LAFOURCADE** and **Anthony PEGARD.**

I find no word to express the gratitude for the support and patience of my family, as I remained away from them during this study period. My wife and son, especially, my daughter having same age as my stay in France. My father, brothers and sisters for their encouragement and support. I always remain in debit of my beloved mother, she has always been a source of inspiration for me.

I am thankful and remained missing my back-home colleague and my teacher, **Kamran ASHRAF,** from University of Veterinary & Animal Sciences Lahore, for constructive discussions and continuous support.

I would show my gratitude to Higher Education Commission of (**HEC**) Pakistan, for providing me final assistance during the course of this study.

I am highly thankful to all the French community, inspired for their open hearted welcoming attitude always with smiling faces, preserving their culture and heritage in such a beautiful manner. Its really ‘un pays merveilleux’.

In the end, of course, for any errors or inadequacies that may remain in this work, the responsibility is solely my own.

Detail of articles resulting from this thesis

Article I

Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P (2012) Carnivore diet analysis based next generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology*, in press**,** doi: 10.1111/j.1365-294X.2011.05424.x.

Article II

Shehzad W, Nawaz MA, Pompanon F, Coissac E, Riaz T, Shah SA, Taberlet P. Jungle without prey: livestock sustains population of an endangered cat in Pakistan. *Oryx*, **submitted**.

Article III

Shehzad W, McCarthy TM, Pompanon F, Purevjav L, Coissac E, Riaz T, Taberlet P. Prey preference of snow leopard (*Panthera uncia*) in South Gobi, Mongolia. *PLoS ONE*, **in revision**.

Other publications

Article IV

Ficetola G, Coissac E, Zundel S, Riaz T, Shehzad W, Bessiere J, Taberlet P, Pompanon F (2010) An *in silico* approach for the evaluation of DNA barcodes. *BMC Genomics*, **11**, 434.

Article V

Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers fom whole genome sequence analysis. *Nucleic Acids Research,* **39**, 11.

Article VI

Taberlet P, Prud'homme S, Campione E, Roy J, Miquel C, Shehzad W, Gielly L, Rioux D, Melo

de Lima C, Pompanon F, Coissac E (2012) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, in press, doi: 10.1111/j.1365-294X.2011.05317.x

Abstract

Information on food webs is central to understand ecosystem functioning. It also provides information of ecosystem stability by evaluating the resource availability and use. Obtaining information on the diet can be critical especially when dealing with elusive carnivores, which are difficult to observe. However, these large carnivores are keystone species that influence the ecosystem through trophic cascades and maintain biodiversity. Thus, precise knowledge of their diet is a prerequisite for designing conservation strategies of these endangered species. Direct and indirect monitoring as well as invasive and non-invasive approaches that have been used to study the diet are either biased or have a low resolution. The DNA-based analysis of feces is an alternative method that may provide better information. It can be implemented through a metabarcoding approach, which is the simultaneous identification of multiple species from a single environmental sample containing degraded DNA by using Next Generation Sequencing. In this case, the use of universal primers for vertebrates amplifying all potential prey also amplifies the predator DNA when it belongs to a close taxon (e.g. mammals). Thus, the PCR products obtained from feces extracts will mainly consist of predator sequences and may not represent the full diet. The use of oligonucleotides specifically blocking the amplification of the predator DNA may overcome this problem. We have developed such a method based on the concomitant use of a universal primer pair (*12SV5*, amplifying all vertebrates) and blocking oligonucleotides for identifying the prey DNA fragments from predators feces. Even if the method developed is not quantitative, it is robust and adequate for studying predator with a very large dietary range and has a better resolution than traditional methods for identifying prey at the genus or species level. This methodology has been applied to characterize the highly eclectic diet (mammals, birds, amphibians and fishes) of two Northern-Pakistani populations of leopard cat (*Prionailurus bengalensis*). With the same approach, we demonstrated the importance of the Human-leopard conflict in Pakistan, due to the almost exclusive consumption of domestic animals by the common leopard (*Panthera pardus*). We could also highlight relevant conservation issues for the highly endangered and cryptic snow leopard (*Panthera uncia*), based on the fact that it mainly fed on wild ungulates.

Résumé

La caractérisation des réseaux trophiques est nécessaire pour comprendre le fonctionnement des écosystèmes et les mécanismes impliqués dans leur stabilité. Il est parfois difficile de déterminer les régimes alimentaires notamment pour des espèces discrètes et difficiles à observer comme les grands carnivores. Cependant, ces espèces jouent un rôle clé dans les écosystèmes dont elles influencent le fonctionnement et la biodiversité. Ainsi, connaitre le régime alimentaire des grands prédateurs avec précision est essentiel pour établir des stratégies de conservation. Diverses méthodes basées sur le monitoring, l'analyse d'échantillons invasifs ou non ont été utilisées pour étudier les régimes alimentaires. Elles sont généralement biaisées ou peu résolutives. Les méthodes basées sur l'identification des fragments d'ADN dans les fèces ont le potentiel de fournir une meilleure information, notamment dans le cadre d'une approche métabarcoding. Il s'agit de caractériser simultanément l'ensemble des espèces dont l'ADN est présent dans un échantillon environnemental, en utilisant les Nouvelles Techniques de Séquençage. Dans ce cas, les amorces universelles nécessaires pour amplifier toutes les proies potentielles amplifient également l'ADN du prédateur s'il y a proximité taxonomique (par exemple mammifères). Ainsi les produits PCR obtenus à partir des fèces sont essentiellement composés d'ADN du prédateur et ne reflètent pas l'ensemble du régime alimentaire. L'utilisation d'un oligonucléotide de blocage limitant spécifiquement l'amplification de l'ADN du prédateur peut résoudre ce problème. Nous avons développé une méthode de ce type basée sur l'utilisation d'amorces universelles pour les vertébrés (amplifiant la région 12SV5) et d'oligonucléotides de blocage. Bien que non quantitative, cette méthode s'est montrée robuste, adaptée à l'étude de prédateurs à très large spectre de proies, et très résolutive pour identifier les proies au niveau du genre et de l'espèce. Nous l'avons appliquée à l'étude du régime alimentaire du chat léopard (*Prionailurus bengalensis*) qui s'est avéré très diversifié (mammifères, oiseaux, amphibiens et poissons) dans les deux populations du Pakistan étudiées. Avec la même approche, nous avons démontré la réalité du conflit entre l'homme et le léopard commun (*Panthera pardus*) dont le régime est presque exclusivement composé d'animaux domestiques. Enfin, nous avons pu proposer des actions de conservations pertinentes après avoir montré que le régime de la très menacée panthère des neiges (*Panthera uncia*) est principalement composé d'ongulés sauvages.

1. Version abrégée en français

Cette thèse présente les résultats des travaux que j'ai effectués pendant 3 ans (2008-2011) au LECA (Laboratoire d'Ecologie Alpine), à l'Université Joseph Fourier de Grenoble. Le manuscrit comporte 6 chapitres. Après ce chapitre de présentation en français, le **chapitre 2** présente le contexte de cette thèse. Il traite de l'utilisation des méthodes de "DNA barcoding' et des Nouvelles Techniques de Séquençage (NTS) pour l'étude des régimes alimentaires. Il fait aussi le point sur les diverses techniques (traditionnelles, moléculaires) utilisées à ce jour pour décrire les régimes alimentaires de carnivores. Les 3 chapitres suivants sont basés sur des articles scientifiques. Le **chapitre 3** traite plus particulièrement du développement de la méthode. Ces développements et les tests de robustesse de la méthode ont été réalisés sur le chat léopard (*Prionailurus bengalensis*) qui a une alimentation particulièrement éclectique. Le **chapitre 4** décrit le régime du léopard commun (*Panthera pardus*), plus particulièrement d'une population du parc national d'Ayubia au Pakistan. Ce cas est particulièrement intéressant dans la mesure où les résultats apportent des connaissances utiles pour gérer les conflits entre population humaine et faune sauvage (impact des prédateurs sur les animaux d'élevage). Le **chapitre 5** relate l'analyse du régime du léopard des neiges (*Panthera uncia*) en Mongolie (Sud Gobi). L'information obtenue, selon laquelle cette espèce menacée se nourrit principalement d'ongulés sauvages, permet de proposer des stratégies de conservation pertinentes. Enfin, le **chapitre 6** propose une conclusion générale dans laquelle le potentiel et les limitations de la technique développée sont présentés. Les perspectives permettent de présenter des solutions aux limitations actuelles de la méthode, ainsi que des recommandation à l'application de cette approche pour de futures études sur d'autres espèces.

1.1 Introduction

L'identification d'espèces à partir de caractéristiques morphologiques est parfois problématique, notamment dans des échantillons environnementaux dans lesquels les organismes sont dégradés voire absent (Valentini *et al.* 2009b). Il reste cependant possible d'identifier les espèces présentes dans un échantillon environnemental en utilisant un code-barres ADN (DNA barcode). Un barcode *sensu stricto* est une région

d'ADN standardisée permettant d'identifier l'espèce. Au sens large, cette notion concerne toute région d'ADN utilisée pour identifier un niveau taxonomique quel qu'il soit (Valentini *et al.* 2009b). Le barcoding *sensu stricto* (Hebert *et al.* 2003b) doit permettre une identification rapide et fiable des organismes animaux grâce à la séquence de la sous-unité 1 du gène mitochondrial de la cytochrome oxydase c (COI). Le concept dans son approche *sensu lato* est utilisé pour faire du métabarcoding, c'est-à-dire identifier l'ensemble des taxa dont l'ADN est présent dans un échantillon environnemental (eau, sol, fèces, Taberlet *et al*. accepté). Une des applications de l'approche métabarcoding est l'étude des régimes alimentaires qui peut se faire à partir de divers types d'échantillons tels que contenus stomacaux (Vestheim & Jarman 2008) ou fèces (Pegard *et al.* 2009; Valentini *et al.* 2009a; Stech *et al*. 2011; Shehzad *et al*. sous presse). Selon la question posée il est possible d'utiliser des couples d'amorces spécifiques d'un groupe de taxons (Deagle *et al*. 2009; Zeale *et al*. 2011) ou universelles (Taberlet *et al*. 2007; Riaz *et al*. 2011; Bienert *et al*. sous presse). L'utilisation de "mini-barcodes" doit permettre de travailler sur des échantillons environnementaux dont l'ADN est dégradé, les fragments amplifiés étant très courts, comme par exemple pour la boucle P6 de l'intron du *trn*L (UAA) ADN chloroplastique (10-150 pb; Taberlet *et al*. 2007). De même, Riaz *et al*. (2011) ont propose un mini-barcode (~ 100 bp) ciblant le gène mitochondrial de l'ARNr 12S. L'approche métabarcoding est particulièrement prometteuse pour étudier le régime alimentaire d'animaux cryptiques ou menacés (Valentini *et al.* 2009b; Shehzad *et al*. sous presse). Elle permet de travailler dans des conditions où il est impossible de caractériser morphologiquement des fragments d'organismes (Zeale *et al*. 2011).

C'est l'apparition des NTS (Nouvelles Techniques de Séquençage) qui a rendu possible le développement des approches métabarcoding. Elles permettent de produire des millions/milliards de séquences par expérience alors que les méthodes de séquençage Sanger sont limitées par l'utilisation de plaques 96-puits (Hert *et al.* 2008; Schuster 2008) ou plus rarement 384-puits (Shibata *et al.* 2000; Emrich *et al.* 2002). Les méthodes de séquençage massif sont commercialisées depuis moins de 10 ans et comprennent les technologies 454 séquençage (454 Genome Séquencer, Roche Applied Science; Basel), Solexa (utilisé dans le Illumina, Genome Analyzer; San Diego), SOLiD (Applied Biosystems; Foster City, CA, USA), Polonator (Dover/Harvard) et HeliScope Single Molecule Sequencer (Helicos; Cambridge, MA, USA). Depuis, de gigantesques

projets de génomiques exploitant ces NTS ont été mis en œuvre (Shendure & Ji 2008). Les NTS ont révolutionné les études de régimes alimentaires à partir de fèces car elles permettent de séquencer un grand nombre d'amplicons obtenus à partir d'un échantillon sans nécessité de les séparer préalablement par clonage. De plus les courtes séquences produites par les NTS sont compatibles avec les courts fragments amplifies à partir d'ADN dégradés (Valentini *et al.* 2009b). La disponibilité croissante de plateformes offrant du séquençage haut-débit à prix compétitif permet des études de régime à grande échelle pour des coûts raisonnables (Deagle *et al*. 2009; 2010; Pegard *et al.* 2009; Valentini *et al*. 2009a). Des études récentes basées sur l'utilisation des NTS ont marqué un tournant dans la caractérisation des régimes alimentaires. Valentini *et al.* (2009a) ont proposé une méthode combinant barcoding ADN et NTS pour analyser le régime d'herbivores en utilisant les amorces universelles *g* et *h* amplifiant un court fragment de la boucle P6 de l'intron du *trn*L (UAA) (Taberlet *et al.* 2007). La comparaison avec des méthodes traditionnelles a démontré la supériorité de l'approche métabarcoding (Soininen *et al*. 2009). Pegard *et al.* (2009) ont montré clairement que l'approche pyroséquençage est beaucoup plus performante que l'approche clonage.

La connaissance du régime alimentaire est un pré requis à la compréhension du fonctionnement des écosystèmes. Cette information permet définir les ressources disponibles et d'évaluer leur consommation dans l'écosystème (Mills 1992). Obtenir ces données est un réel challenge pour les espèces de carnivores élusives ou fortement menacées. Ces carnivores sont en effet des espèces clé pour un écosystème; elles ont une forte influence sur son fonctionnement même si elles représentent une part modérée de sa biomasse (Mills *et al.* 1993; Power *et al.* 1996).

De nombreuses approches aux potentialités et limitations différentes ont été utilisées pour l'étude des régimes alimentaires (Shrestha & Wegge 2006). L'observation directe ou indirecte par vidéo, l'utilisation de données de radio tracking (suivi télémétrique) permettant de découvrir des restes de proies sur le trajet du prédateur, la collecte de témoignages peuvent être utilisées. Quand elle est possible, l'observation donne des informations très précise comprenant aussi l'âge, la taille, le sexe de la proie. Mais ces techniques ne sont pas opérationnelles pour étudier de très nombreux individus. Différentes méthodes invasives ont aussi été largement utilisées, comme l'analyse de contenus intestinaux ou stomacaux. Mais elles ne sont pas applicables sur des grands carnivores ou animaux menacés d'extinction car elles nécessitent de tuer l'animal ou au

moins de l'immobiliser pour purger son estomac. L'analyse brute ou histologique de contenus stomacaux ou intestinaux (Prado *et al*. 2010; Soininen *et al.* 2009; Balestrieri *et al.* 2011) a l'avantage de permettre l'identification des proies consommées quel que soit leur niveau de digestibilité (Hyslop 1980). Ce type d'échantillon a aussi été utilisé pour caractériser les espèces de proies à partir l'électrophorèse de protéines ou de différentes techniques immunologiques (test de précipitation de protéines, western blot, ELISA). Les principales limitations de ces techniques immunologiques sont les faux positifs et est le partage de déterminants antigéniques par les différentes espèces de proies possibles (Feller *et al.* 1985). De plus, bien que les tests ELISA permettent la détection de faibles quantités de protéines d'une proie ils ne sont pas quantitatifs sur ce type d'échantillons stomacaux (Hoyt *et al*. 2000). Le développement d'anticorps mono- ou poly-clonaux est également un processus lourd à mettre en œuvre (Sheppard & Harwood 2005).

L'identification de proies par étude des fèces a été largement répandue, notamment de par la nature non-invasive de l'échantillon permettant d'étudier des espèces menacées ou difficilement observables. L'examen brut de restes non digérés (os, coquilles, poil) ou l'observation microscopique des phanères sont possibles mais fournissent une information limitée pour une grande quantité de travail (Oli 1993; Hall-Aspland & Roger 2007). Le biais dû à une digestibilité différentielle des proies peut être évité par l'utilisation de méthodes comme la chromatographie en couche mince, la caractérisation des isotopes stables ou la Spectrographie Proche Infra-Rouge (SPIR). La Caractérisation des isotopes stables par exemple sur des tissus du prédateur donne des informations à long terme sur son alimentation mais ne permet pas l'identification des espèces ingérées. Après calibration la méthode SPIR est rapide et peu coûteuse, cependant sa résolution reste faible et peut générer des faux positifs.

L'analyse moléculaire des fèces (Höss *et al.* 1992; Kohn et Wayne 1997; Waits et Paetkau 2005) fournit des données plus fiables et plus précises. Cette approche est particulièrement efficace pour les espèces difficilement observables et menacées. Un problème important est lié à la présence de l'ADN du prédateur qui rend difficile l'utilisation de systèmes d'amorces universels permettant d'amplifier l'ADN de toutes les proies possibles sans connaissance *a priori* du régime alimentaire. Si les approches universelles sont efficaces pour les herbivores (Valentini *et al.* 2009a), elles représentent un réel challenge pour les carnivores, l'ADN du prédateur étant co-amplifié avec celui

des proies (Jarman *et al*. 2006) et pouvant devenir dominant dans le produit PCR (Green & Minz 2005, Jarman *et al.* 2004, 2006) masquant la présence de proies.

Plusieurs méthodes ont été proposées pour permettre l'amplification des ADN minoritaires de proies. L'utilisation d'amorces spécifiques d'une espèce ou d'un groupe d'espèces excluant le prédateur (Vestheim *et al*. 2005; Deagle *et al*. 2006; Juen & Traugott 2007; King *et al.* 2010; Zeale *et al*. 2011) n'est pas adéquate lorsque que le spectre de proies est très large. De plus la définition d'amorces spécifiques n'est pas toujours facile si les régions étudiées sont peu variables entre proies et prédateur. Des enzymes de restriction coupant spécifiquement l'ADN du prédateur ont été utilisées avec des amorces universelles pour montrer la diversité du régime alimentaire dans plusieurs études (e.g. Blankenship & Yayanos 2005; Suzuki *et al*. 2006, 2008; Dunshea 2009). Mais cette méthode n'est efficace que si l'ADN du prédateur est suffisamment différent de celui des proies, et nécessite une connaissance *a priori* du spectre de proies.

Des oligonucléotides présentant des spécificités de leur mode de synthèse (e.g. PNA, LNA, DPO et C3 espaceur) ont été utilisés pour bloquer l'amplification de fragments non cibles même en excès dans l'échantillon. Ces oligonucléotides sont définis pour être spécifiques du fragment non souhaité et s'hybrident préférentiellement avec lui et empêchent son élongation par la polymérase lors de la PCR. Les autres fragments, même minoritaires, sont alors amplifiés. L'utilisation conjointe de ce type d'oligonucléotides avec des amorces universelles a permis plusieurs études (c-à-d Vestheim & Jarman 2008; Deagle 2009, 2010; Hu *et al.* 2009; Chow *et al*. 2011). Cette approche de blocage a de nombreux avantages. Par exemple elle ne nécessite pas d'avoir une information *a priori* sur la nature des proies consommées. Elle peut être mis en œuvre avec un surcout modéré dépendant du type d'oligonucléotide utilisé. Par exemple, les oligonucléotides PNA sont 6 fois plus chers que ceux présentant un espaceur C3. Un problème concernant l'ensemble des méthodes moléculaires est qu'elles ne permettent pas de détecter les cas de cannibalisme.

1.2 Nouvelles Techniques de Séquençage pour l'analyse du régime alimentaire d'une carnivore: application au chat léopard au Pakistan

Chapitre issu de l'article: Carnivore diet analysis based next generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan

Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P. (article en révision dans Molecular Ecology).

L'article présenté dans ce chapitre présente la méthode alternative à l'étude des restes non digérés, pour déterminer le régime alimentaire des carnivores à partir de fèces. Cette méthode basée sur l'utilisation des NTS est appliquée au chat léopard par l'étude de 2 populations issues de régions du Nord Pakistan écologiquement contrastées ANP et CGNP (Ayubia National Parc et Chitral Gol National Parc). Bien que le chat léopard soit relativement commun dans une grande partie de l'Asie, on dispose de peu d'informations sur ce prédateur peu visible.

L'étude des régimes alimentaires de carnivores est basée sur l'utilisation d'un barcode amplifié avec les amorces *12SV5F*/*12SV5R* définies à l'aide du programme *ecoPrimers* (Riaz *et al.* 2011). *ecoPrimers* détermine des barcodes et leurs amorces associées en scannant des bases de données génomiques. Il retient les solutions optimisant la couverture taxonomique et le pouvoir discriminant du barcode, prenant en compte des contraintes prédéfinies telles que la taille du marqueur ou les taxons ciblés. Le marqueur *12SV5F*/*12SV5R* est un barcode universel pour les vertébrés défini à partir des bases de génomes mitochondriaux complets disponibles à ce jour. Les amorces amplifient un fragment d'environ 100 pb de la boucle V5 du gène mitochondrial 12S. Malgré une faible taille, qui permet l'amplification à partir d'ADN issu de fèces, ce barcode à un fort pouvoir discriminant. Riaz *et al.* (2011) ont montré qu'il identifie 77 % des espèces et 89 % des genres de vertébrés. Etant donné que les systèmes universels tels que celui-ci amplifient à la fois l'ADN des proies et celui du prédateur, notre travail a eu pour objectif de développer une méthode bloquant l'amplification du prédateur (ici le chat léopard). Sur le fragment cible, nous avons défini un oligonucléotide chevauchant de 6 bases la séquence de l'amorce anti-sens et s'étendant jusqu'à une

région spécifique du chat léopard (Fig. 2.2e). Nous avons testé l'effet de cet oligonucléotide avec des PCR utilisant le couple d'amorces *12SV5F/12SV5R* en absence ou présence de l'oligonucléotide de blocage *PrioB* (20 fois plus concentré que les amorces). Le séquençage NTS des produits PCR a montré l'efficacité de dispositif (voir Fig. 1.1).

L'analyse de 22 échantillons fécaux de la population ANP et de 16 de la population CGNP a révélé 18 types de proies dont 8 mammifères (dans 87 % des fèces), 8 oiseaux (34 %), un amphibien (16 %) et un poisson (3 %) (Fig. 1.2). Sept proies n'ont été rencontrées que dans un échantillon. Quatre de ces 18 proies (*Apodemus rusiges*, *Dryomys nitedula*, *Eoglaucomys fimbriatus*, *Lepus capensis*) sont des espèces vulnérables (Sheikh & Malour 2004). Le chat léopard qui est répandu au Pakistan est donc un danger potentiel pour ces espèces que l'on retrouve dans 44,7 % (en fréquence cumulée) des fèces. Une étude plus approfondie doit avoir lieu pour évaluer l'importance de ce danger et proposer des stratégies de gestion appropriées prenant en compte à la fois les prédateurs et les proies. Une plus grade diversité des proies a été observée en ANP, reflétant probablement la plus grande productivité et diversité des forêts tempérées de ce parc. Nos résultats confirment le large spectre de proies du chat léopard, composé surtout de rongeurs (principalement Muridae). La comparaison de nos résultats à ceux d'études menées plus au Sud de son aire de répartition montre une tendance à la consommation de proies plus grosses au Pakistan. La méthode ADN universelle développée ici est simple à mettre en œuvre, particulièrement robuste et ouvre la voie à des analyses à grande échelle. Elle apporte donc un plus par rapport aux méthodes traditionnelles et peut être étendue à d'autres carnivores en modifiant les oligonucléotides de blocage de l'ADN du prédateur.

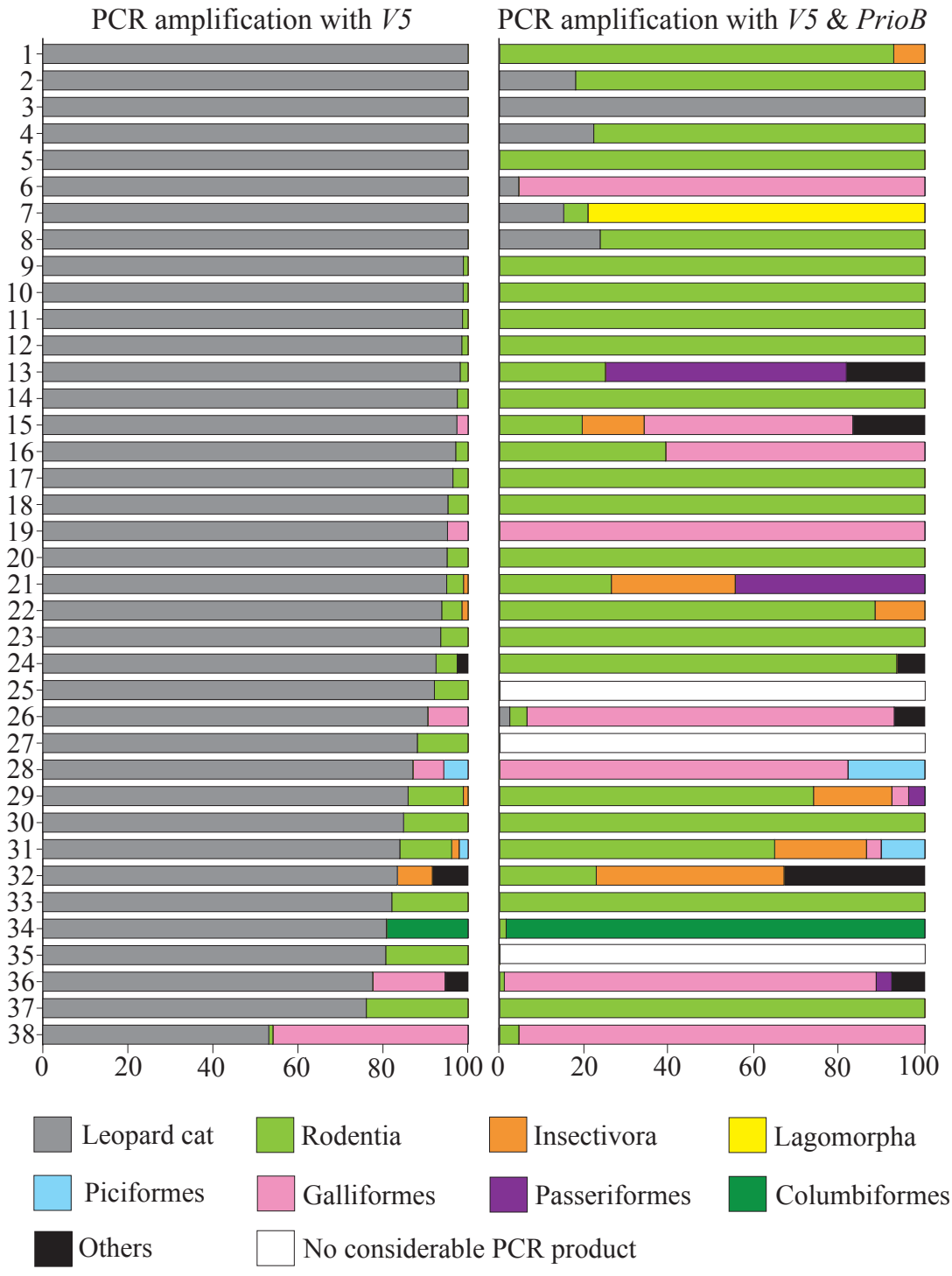
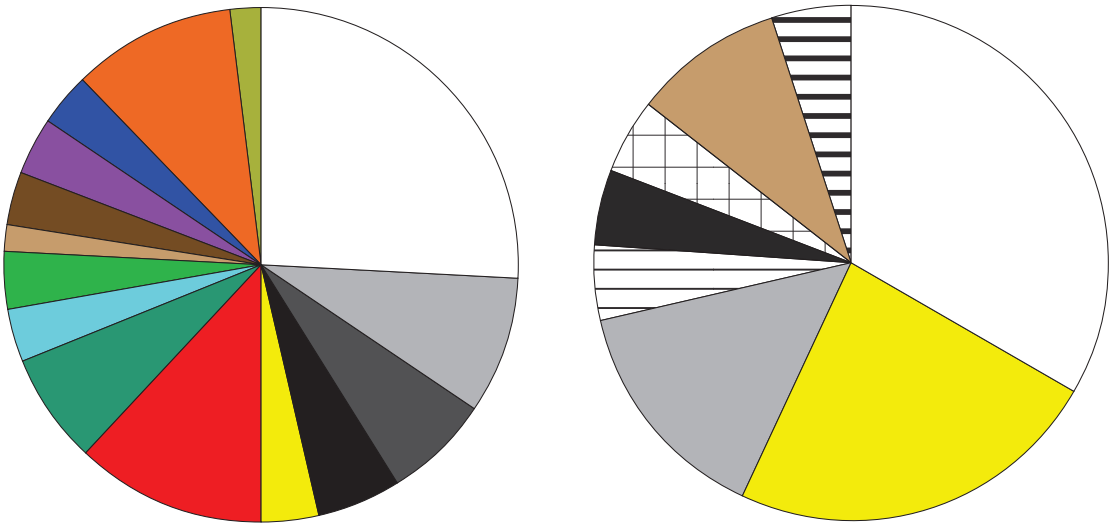


Fig. 1.1 Comparaison de l'amplification avec ou sans oligonucleotide de blocage (pourcentage). Les taxons des proies sont donnés au niveau de l'ordre, les poissons et amphibiens sont regrpupés dans la catégorie "others".

a) Ayubia National Park b) Chitral Gol National Park



Mammals

House rat (*Rattus rattus*)

Himalayan wood mouse (*Apodemus rusiges*)

Murree vole (*Hyperacrius wynnei*)

House mouse (*Mus musculus*)

Kashmir flying squirrel (*Eoglaucomys fimbriatus*)

Forest dormouse (*Dryomys nitedula*)

Asiatic white toothed shrew (*Crocidura pullata*)

Cape hare (*Lepus capensis*)

Birds

Kalij pheasant (*Lophura leucomelanos*)

Chicken (*Gallus gallus*)

Koklass pheasant (*Pucrasia macrolopha*)

Chukar partridge (*Alectoris chukar*)

Babbler (Timaliidae)

Jungle crow (*Corvus macrorhynchos*)

Woodpecker (*Dendrocopos* sp.)

Rock pigeon (*Columba livia*)

Amphibian

Murree hill frog (*Paa vicina*)

Fish

Cat fish (Siluriformes)

Fig. 1.2 Comparaison du régime alimentaire du chat léopard entre les Parcs Nationaux d'Ayubia et du Chitral Gol

1.3 Une jungle sans proie : maintien d'une population menacée de panthère par exploitation des ressources domestiques

Chapitre issu de l'article : Jungle without prey: livestock sustains population of an endangered cat in Pakistan

Shehzad W, Nawaz MA, Pompanon F, Purevjav L, Coissac E, Riaz T, Shah SA, Taberlet P. (submitted in Oryx).

Le conflit entre humains et grands carnivores est l'un des grands défis de la gestion des populations de grands carnivores. Les principales préoccupations sont (i) l'augmentation de la population humaine au cours de la dernière décennie, (ii) la dégradation et la fragmentation des forêts amenant à la perte d'habitat pour les populations de carnivores, et (iii) la prédation du bétail dû au manque de proies naturelles, conduisant aux abattages de carnivores sauvages commis en représailles. Une véritable estimation des espèces de proies consommées par un prédateur est un repère nécessaire pour évaluer l'ampleur de ce conflit. Nous avons utilisé des analyses récemment développées du régime alimentaire à partir de l'ADN pour étudier le profil des proies de la population de léopard commun du Parc National Ayubia (Ayubia National Park, ANP) au Pakistan. Les résultats de cette étude ont clairement démontré que le léopard commun est un prédateur généraliste qui se nourrit principalement d'animaux domestiques. D'après les fréquences d'apparition des proies dans 57 échantillons de matières fécales, la chèvre domestique prédominait dans le régime alimentaire (64,9 %), suivi du chien (17,5 %) et de la vache (12,3 %). Les animaux domestiques (chèvres, chiens, vaches, buffles, chevaux et moutons) apparaissent dans 54 des 57 matières fécales, ce qui correspond à une fréquence d'apparition de 0,95. Une stratégie en deux étapes devraient être adoptée pour atténuer le conflit, premièrement en introduisant des compensations financières aux éleveurs à l'intérieur et autour de la zone protégée, ensuite en permettant de rendre à nouveau disponible des proies sauvages. Nous espérons que les résultats de cette étude contribueront à la survie des populations de léopard commun au Pakistan.

1.4 Préférence alimentaire du léopard des neiges dans le Sud du Gobi, Mongolie

Chapitre issu de l'article : Prey preference of snow leopard (*Panthera uncia*) in South Gobi, Mongolia.

Shehzad W, McCarthy TM, Pompanon F, Coissac E, Riaz T, Taberlet P. (Submitted in PLoS ONE).

L'information sur le régime alimentaire des grands carnivores qui sont discrets et qui habitent des terrains difficiles d'accès est importante pour bien concevoir des stratégies de conservation. Partout sur l'ensemble du territoire du léopard des neiges, la prédation sur le bétail et les abattages commis en représailles sur les prédateurs sont devenus un sérieux problème. D'ailleurs, plusieurs études ont été réalisées sur l'écologie alimentaire du léopard des neiges en utilisant des approches classiques. Ces techniques ont d'inhérentes limites pour identifier correctement aussi bien les excréments de léopard des neiges que les taxons des proies. Afin d'examiner dans le régime alimentaire du léopard des neiges la fréquence d'apparition des proies de bétail et de l'argali presque menacé, nous avons utilisé une approche récemment développée sur le régime alimentaire à partir de l'ADN, et l'avons appliquée à une population de léopards des neiges des montagnes Tost, dans le Sud de Gobi en Mongolie. Après avoir extrait l'ADN à partir de fèces, utilisé des amorces universelles pour les vertébrés et un oligonucléotide de blocage spécifique de l'ADN du léopard des neiges, une région de ~100 paires de bases du gène mitochondrial de l'ARN 12S a été amplifié. Les amplicons ont ensuite été séquencés sur une plate-forme de séquençage de nouvelle génération. Nous avons observé un total de cinq proies différentes dans 81 échantillons de fèces. Le bouquetin de Sibérie prédomine dans le régime alimentaire (dans 70,4% des fèces), suivi de la chèvre domestique (17,3%) et du mouton argali (8,6%). Les grands ongulés composent la majeure partie de l'alimentation (dans 98,8% des fèces), en incluant les ongulés sauvages (79%) et le bétail domestique (19,7%). Les résultats de cette étude vont permettre de comprendre l'écologie alimentaire du léopard des neiges, aussi bien pour répondre aux questions de conservation qu'aux questions de gestion appliquées à ce félidé sauvage.

1.5 Conclusions et perspectives

La combinaison d'oligonucléotides de blocage avec des amorces universelles permettant d'amplifier l'ADN de tous les vertébrés à partir d'un extrait de fèces, suivi par séquençage du produit PCR à l'aide de NTS ouvre de large perspectives dans l'étude des régimes alimentaires de carnivores. L'oligonucléotide de blocage permet de réduire l'amplification de l'ADN du prédateur (par exemple à 2.2 % des séquences produites dans le cas du chat léopard), et d'amplifier l'ADN de proies même lorsqu'il est minoritaire dans l'extrait de départ. Cette technique basée sur l'analyse d'échantillons non invasifs est adaptée à l'étude d'espèces menacées et permet de caractériser des régimes alimentaires extrêmement diversifiés (cas du chat léopard). Nous avons montré que cette technique peut facilement être adaptée à différents types de carnivores par modification de l'oligonucléotide de blocage (léopard commun et léopard des neiges mais également loup, non présenté ici). L'efficacité de cette méthode repose en grande partie sur l'utilisation des NTS, permettant de séquencer en une expérimentation un grand nombre de produits PCR, pour peu qu'ils aient été générés avec des amorces marquées (le tag étant un court oligonucléotide spécifique de l'échantillon ajouté aux amorces et permettant de trier les séquences produites lors de l'analyse). Ainsi cette méthode est particulièrement adaptée aux études à grandes échelles combinant des centaines d'échantillons.

Ce système est robuste et a une meilleure résolution que les méthodes traditionnelles dans l'identification des proies. Nous l'avons démontré en comparant l'approche ADN avec l'examen microscopique des poils dans des fèces de loup (Shehzad *et al.* en prép; article non présenté dans ce manuscrit). La technique ADN lève de nombreuses ambigüités concernant des espèces proches de proies difficilement distinguables par l'examen microscopique. Il est également possible de combiner l'analyse ADN du régime alimentaire avec un génotypage permettant l'identification individuelle et le sexage. On peut ainsi réaliser une analyse complète de la population sans suivi direct sur le terrain, à partir d'échantillons non invasifs.

La méthode développée ne nécessite aucune information *a priori* sur les proies, ce qui n'est pas le cas lorsque l'on utilise de systèmes d'amorces spécifiques (d'une espèce ou d'un groupe d'espèces), ou que l'on recherche des sites de restrictions absents des

séquences de proies. L'utilisation d'amorces de blocage est également peu coûteuse en comparaison des autres types d'oligonucléotides modifiés (par exemple six fois moins cher que les PNA; Chow *et al*. 2011).

Une question encore débattue est l'aspect quantitatif des analyses ADN de régimes alimentaires, à savoir si le nombre de séquences obtenues pour différentes espèces d'un échantillon est proportionnel aux proies consommées. Cet aspect a été abordé dans plusieurs études récentes (c-à-d Valentini *et al.* 2009a; Soininen *et al.* 2009; Deagle *et al.* 2009, 2010). Le nombre de séquences ne peut pas être interprété de façon quantitative pour plusieurs raisons. L'utilisation d'amorces universelles sur des mélanges prédéterminés a pu montrer un biais d'amplification en faveur de certaines espèces (Polz & Cavanaugh 1998). Ces biais sont notamment observés lorsque la séquence de certaines espèces diffère de celle des amorces, favorisant l'amplification des fragments d'ADN où des amorces s'hybrident parfaitement. Dans notre approche les régions ciblées par les amorces universelles *V5F* et *V5R* sont très fortement conservées (article I) limitant les biais de PCR. Deagle *et al*. (2010) ont suggéré que des variations en densité de mitochondries selon les échantillons peuvent également faire varier le nombre de séquences obtenues. Dans nos études, nous avons donc évité les interprétations quantitatives, ne travaillant qu'en présence/absence de proies par fèces.

Au cours de notre étude, nous avons dû également faire face à limitations générales à tous types d'études ADN du régime alimentaire. La caractérisation ADN de fèces ou de contenus stomacaux donne une vision instantanée de l'alimentation qui peut différer d'un comportement à long terme. Une récolte suivie des fèces n'est pas toujours possible pour des espèces très discrètes. Dans ce cas l'étude des signatures isotopiques qui produit des informations à long terme à partir de poils ou de tissus du prédateur est intéressante. Elle est complémentaire des analyse ADN, car seule, elle ne permet pas une différenciation des proies au niveau spécifique. Une autre limitation de données ADN est l'impossibilité de différencier la prédation de la saprophagie. On peut alors surestimer le niveau de prédation dans la caractérisation d'un réseau trophique. Des méthodes immunologiques permettent de distinguer la consommation de proies vivantes de celle de cadavres (Clader *et al.* 2005) et offrent ainsi une approche complémentaire des techniques que nous avons développées. Les approches ADN ne permettent pas non plus d'estimer l'âge des proies. Quelques tentatives ont été entreprises pour déterminer l'âge d'individus d'espèces difficilement étudiables (voir le revue de Dunshea *et al*.

2011). L'utilisation de la dégradation des télomères comme marqueur de l'âge ne semble cependant pas une méthode fiable, et seule l'observation de la proie ou de ses restes peut aujourd'hui donner une information précise. La méthode que nous avons développée ne permet pas non plus de détecter les cas de cannibalisme, par exemple dans le cas de la prédation de jeunes, puisque leur ADN ne peut être distingué de celui du prédateur. Plus généralement, les amorces utilisées (*12SV5F* et *12SV5R*) ne permettent de caractériser que des vertébrés alors que de nombreux carnivores ont un régime plus diversifié incluant des invertébrés et même des plantes. Par exemple le blaireau (*Meles meles*) a un large spectre de proie incluant amphibiens, reptiles, oiseaux et petits mammifères, mais aussi vers de terre, insectes et racines (Revilla & Palomares 2002). Dans le cas de régimes omnivores, il est possible de combiner différentes amorces à celles ciblant les vertébrés, pour caractériser les plantes (e.g. Taberlet *et al.* 2007; Valentini *et al.* 2009a) ou les vers de terre (Bienert *et al*. sous presse). D'autres problèmes reposent sur la qualité et l'exhaustivité des bases de références utilisées pour identifier les taxons. Par exemple la détection de proies endémiques de la zone d'étude peut être problématique si elle n'est pas présente dans les bases publiques comme Genbank/EMBL/DDBJ (voir discussion de l'article I). Ainsi, nous recommandons de construire des bases de références locales pour une caractérisation précise du régime alimentaire.

Comme toutes techniques de séquençage les Nouvelles Techniques de Séquençage sont sources d'erreurs par incorporation de mauvaises bases lors de la synthèse. La technique Illumina est plus sensible aux substitutions qu'aux insertions délétions, alors que la technique 454 à plus de problème dans le séquençage d'homopolymères. Lors du séquençage de produits PCR elles permettent également de révéler les erreurs produites lors de l'amplification. Ces erreurs sont d'autant plus probables lors de l'analyse d'échantillons environnementaux contenant les ADN de plusieurs espèces. Dans ce cas, il est possible de produire des séquences chimères qui se forment quand un fragment incomplètement synthétisé (ou plus court) à un moment de la PCR joue le rôle d'amorce sur un fragment d'une autre espèce au cycle suivant. Nous avons observé beaucoup de séquences chimériques dans nos résultats, bien qu'elles soient en proportion relativement faible par rapport aux séquences exactes. Ce type d'erreur est gênant pour l'identification des espèces, mais les chimères ont généralement de faibles scores d'identité dans les bases de référence.

Les écologistes moléculaires se tournent maintenant de plus en plus vers l'étude de l'ADN des régimes alimentaires. Malgré des contraintes techniques, qui doivent être levées, telles que la sélection de métabarcodes et la quantification de proies dans un échantillon, ces nouvelles techniques ont un fabuleux potentiel. Elles fournissent des données précises et moins biaisées que celles obtenues par d'autres techniques plus classiques. Des études récentes ont permis de caractériser des régimes herbivores (Valentini *et al*. 2009a), carnivores (notre étude) et omnivores (Alfstrom *et al*. en prép.). Alternativement, alors que l'utilisation de PCR basées sur les métabarcodes est toujours problématique, nous suggérons l'application d'une approche de «sonde de capture»**,** en utilisant les régions conservées d'oligonucléotides ciblés. Ces sondes de capture permettent de récupérer des fragments cible qui seront séquencés en utilisant les NTS pour identifier les différents taxons présents dans un échantillon environnemental. La probabilité de trouver une sonde conservée est plus élevée que celle de trouver des paires d'amorces. En outre, de nombreuses sondes, ciblant différents groupes taxonomiques peuvent être multiplexées, en une seule expérimentation, pour révéler tous les taxons (voir Taberlet *et al.* article soumis).

L'analyse de régime basée sur la présence/absence de proies est très informative. Cependant, comme nous l'avons souligné, il est difficile d'obtenir une information quantitative (proportions des proies consommées). De même, l'estimation du nombre d'individus consommés d'une même espèce n'est pas estimable. Sans cet aspect quantitatif, les études ADN fourniront une information incomplète. Deagle *et al*. (2010) ont abordé l'étude de ces aspects quantitatifs et ont montré les limites des techniques basées sur l'ADN. Dans cette optique, des études pilotes sont en cours dans notre laboratoire en utilisant des multiplexes de barcodes afin d'étudier des régimes omnivores. A l'avenir, nous suggérons également d'essayer une autre approche pour trouver les différentes proportions des taxons des différentes proies présentes dans un échantillon environnemental en utilisant le séquençage direct des extraits d'ADN de l'environnement. Les plate-formes NTS ont la capacité de produire des millions de séquences appartenant à différents taxons. Même si seulement une petite partie de ces séquences pourrait être identifiée en recherchant dans l'ensemble des bases de données, suffisamment d'informations serait disponible pour décrire la diversité de l'échantillon. Par ailleurs, une approche telle que le shotgun pourrait avoir l'avantage de permettre

d'identifier les proportions relatives entre les différents types d'ADN provenant de différents groupes taxonomiques. (Taberlet *et al*. article soumis).

Les résultats présentés dans ce manuscrit sont le reflet des régimes étudiés sur une saison. Cependant, nos recommandons que les études écologiques nécessitant ce type de données soient menées sur le long terme. C'est cette dimension temporelle qui doit permettre une vision précise des relations prédateur/proie ainsi et de leurs variations.

2. Introduction

2.1 DNA barcode –a reliable tool for species identification

Accurate species identification is fundamental to recognize and describe biodiversity. Traditional identification based upon morphological characteristics with identification keys, is not always evident (Hajibabaei *et al.* 2007). For example some organisms like insects have entirely different morphological features from larval to adulthood (Jinbo *et al*. 2011). Moreover, morphological basis identification requires a lot of expertise and time, and is still prone to many errors (Hebert & Gregory 2005; Hajibabaei *et al.* 2011). Recent advances have proposed the feasibility of introducing an identification system based on the sequence variation of small DNA fragments (Hebert *et al*. 2003a,b; Valentini *et al*. 2009b). A barcode is a standardized region of DNA designated for reliable identification either to the species level (*sensu stricto*; Hebert *et al.* 2003b) or to any taxonomical level (*sensu lato;* Valentini *et al.* 2009b).

The main objective of this approach (DNA barcoding) is to establish a large-scale reference sequence database against which unknown samples can be compared for identification. This concept was first introduced by Arnot *et al.* (1993) and did not gained much acceptability among scientific community. Later, Hebert *et al.* (2003b) suggested a quick and accurate identification system from biological samples, by amplifying the mitochondrial *cytochrome c oxidase* subunit I (*COI*) gene and referred this approach as 'DNA barcoding'. Afterwards, the Consortium for the Barcode of Life (CBoL, http://barcoding.si.edu) was proposed with the aim of developing standard protocols for DNA barcoding and to create a comprehensive DNA barcode library. More recently, in October 2010, this project has entered into a new phase with the launch of the International Barcode of Life project (iBoL; http://ibol.org/), relying on the alliance of 27 countries that aim to establish an automated identification system based on a DNA barcode library of all eukaryotes.

An ideal DNA barcode should meet the following criteria as suggested by Taberlet *et al.* (2007). First, it must be highly conserved within the same species, but variable among all species to identify. Second, it should be standardized i.e. the same DNA region has the ability to identify species from various taxonomic groups. Third, the target DNA region should contain enough phylogenetic information to easily

assign an unknown species to its taxonomic group (genus, family, etc.). Fourth, the variable region should be between two highly conserved regions, so that the fragment can be amplified for a broad range of species. Fifth, the target DNA region should be short enough (about 150 bp) for allowing amplification of degraded DNA recovered from environmental samples. Finding a DNA barcode marker, which fulfills all the above criteria, is quite critical. However, for different group of users (e.g. taxonomists versus ecologists) the priority given to the above five criteria may be different. For example, for taxonomists, a high level of variation with sufficient phylogenetic information is required, while ecologist will favor short and robust barcodes.

Advances in bioinformatics have guided the choice of universal barcodes by selecting more resolutive regions with fewer ambiguities (Bellemain *et al.* 2010; Ficetola *et al.* 2010). For example, Riaz *et al.* (2011) developed a program called '*ecoPrimers*' that scans whole genome databases (e.g. mitochondrial genome databases) to find such markers that will be used with no need for *a priori* information about the prey species. This program optimizes two quality indices called barcode coverage and barcode specificity to select the most efficient markers from a set of reference sequences, according to specific experimental constraints such as the length of the marker or the range of target taxa. Barcode coverage evaluates the ability of primers to amplify taxa within the target group. Barcode specificity measures the ability of markers to discriminate among taxa. The recently developed Next Generation Sequencing (NGS) technologies combined with DNA barcoding have revolutionized ecological studies (Valentini *et al.* 2009b), as huge amounts of sequence reads can be generated to characterize the complete biodiversity of a PCR product obtained from environmental samples (Jinbo *et al.* 2011).

DNA barcoding has a tremendous potential for being used in various fields. It can be used in diverse domains of biological sciences including taxonomy, ecology, conservation biology, agriculture, fish management and forestry (Hollingsworth 2007). DNA Barcoding ideally allowed molecular based identification of unknown taxa, where morphological identification can offer only approximate assignments up to the genera /family or order level (Darling & Blum 2007). It presents a direct approach, involving analysis of DNA sequence for each sample.

Species identification from environmental samples is fundamental in ecological studies (Damm *et al.* 2010). The DNA recovered from non-invasive sampling

techniques is usually degraded and fragmented (Taberlet *et al.* 1999). DNA barcoding is an ideal system for studying such samples by proposing markers that can amplify short and degraded DNA sequences. DNA metabarcoding is a new concept dealing with the simultaneous characterization of many taxa present in an environmental sample (e.g. feces or soil), by amplification of its DNA with universal primers and next generation sequencing of the PCR product (Taberlet *et al.* accepted).

DNA barcoding/metabarcoding has the remarkable potential for studying the diet of endangered animals. It offers quite a flexible range of different sampling methods from stomach contents (Vestheim & Jarman 2008) to feces analysis (Pegard *et al.* 2009; Valentini *et al.* 2009a; Stech *et al*. 2011; Shehzad *et al*. in press). Depending on the question addressed, specific barcodes (Deagle *et al*. 2009; Zeale *et al*. 2011) or universal primers (e.g. Taberlet *et al*. 2007; Riaz *et al*. 2011; Bienert *et al*. in press) could be used for prey identification. Recently "mini-barcodes" have been developed, studying especially the diet from highly degraded environmental samples. These mini-barcodes are very short fragments, for example, the 'g-h' part of the *trn*L (UAA) P6 loop of the chloroplast DNA, to study plant taxa (10-150 bp, length of amplified fragment excluding length of primers; Taberlet *et al*. 2007). Similarly, Riaz *et al*. (2011) have proposed a mini-barcode (~ 100 bp) targetting the mitochondrial 12rRNA gene of vertebrates. DNA barcoding is quite promising in study the diet where animals are endangered and difficult to observe directly (Valentini *et al.* 2009b; Shehzad *et al*. in press). It is even able of analyze prey items consumed where identification of morphological features are impossible (Agusti *et al.* 2003; Zeale *et al*. 2011).

2.2 Next generation sequencing systems

Technological advances in sequencing platforms have changed the scenario of many ecological studies including diet analysis (Deagle *et al.* 2009; Valentini *et al.* 2009a). The NGS (Next Generation Sequencing) systems have increased throughput by use of massive parallel DNA sequencing systems (Kircher & Kelso 2010). The sequence of a huge number of different DNA molecules can be determined in a single experiment at a substantially low price (Pegard et *al*. 2009; Schadt et *al*. 2010). These systems allow millions of 'reads' to be produced in a single experiment, in contrast to existing Sanger capillary sequencer, the throughput of which is mostly limited either by a 96-sequencing templates (Hert *et al.* 2008; Schuster 2008) or rarely 384-

sequencing templates (Shibata *et al.* 2000; Emrich *et al.* 2002). In order to study the full complexity of PCR products such as those obtained from environmental samples, the Sanger capillary sequencing is fairly tedious and inefficient, as it requires prior multistep cloning procedures of different amplified DNA fragments. This makes this method not cost effective (Margulies *et al.* 2005; Pegard *et al.* 2009).

NGS with rapidly growing technology refers to several innovative systems. They arrived in the commercial marketplace only over the last decade and comprise the 454 sequencing (used in the 454 Genome Sequencers, Roche Applied Science; Basel), the Solexa technology (used in the Illumina, Genome Analyzer; San Diego), the SOLiD platform (Applied Biosystems; Foster City, CA, USA), the Polonator (Dover/Harvard) and the HeliScope Single Molecule Sequencer technology (Helicos; Cambridge, MA, USA). Since then, tremendously new genome research projects have been undertaken that exploit the awesome throughput advantages of these systems (Shendure & Ji 2008).

In this study we used Illumina sequncing for analyzing the diet. The Genome Analyzer IIx of Illumina works on the principle of sequencing by synthesis. It uses reversible terminator chemistry with four different fluorescently labeled nucleotides. First a library of the sequences that need to be determined must be prepared (Bentley *et al*. 2008; Metzker 2010): two different adapters are ligated at the each end of a DNA molecule, which is an amplicon from a PCR product in our case. Then each DNA molecule is covalently attached to a glass surface by one of the adapters (Fig. 2.1). This library is then amplified (solid state amplification) using unlabelled primers and a polymerase. Several cycles of amplifications are undertaken to create a dense cluster of DNA template surrounded by adapters at both ends. This step is called solid-phase bridge amplification. Then, sequencing starts with the addition of fluorescently labeled dNTPs terminators and a DNA polymerase. As the complementary base has been incorporated at first position, the polymerization is stopped for a short time and unincorporated nucleotides and other reagents are then washed away. A laser is used to excite and read this bound fluorescent nucleotide and the signal is recorded. The fluorescent label and the terminating group attached to the incorporated base are then removed, allowing for further extension of DNA fragment (Ansorge 2009; Kicher & Kelso 2010; Metzker 2010). Illumina has recently introduced a technical update by reverse strand sequencing of each DNA molecule called “paired-end reads”.

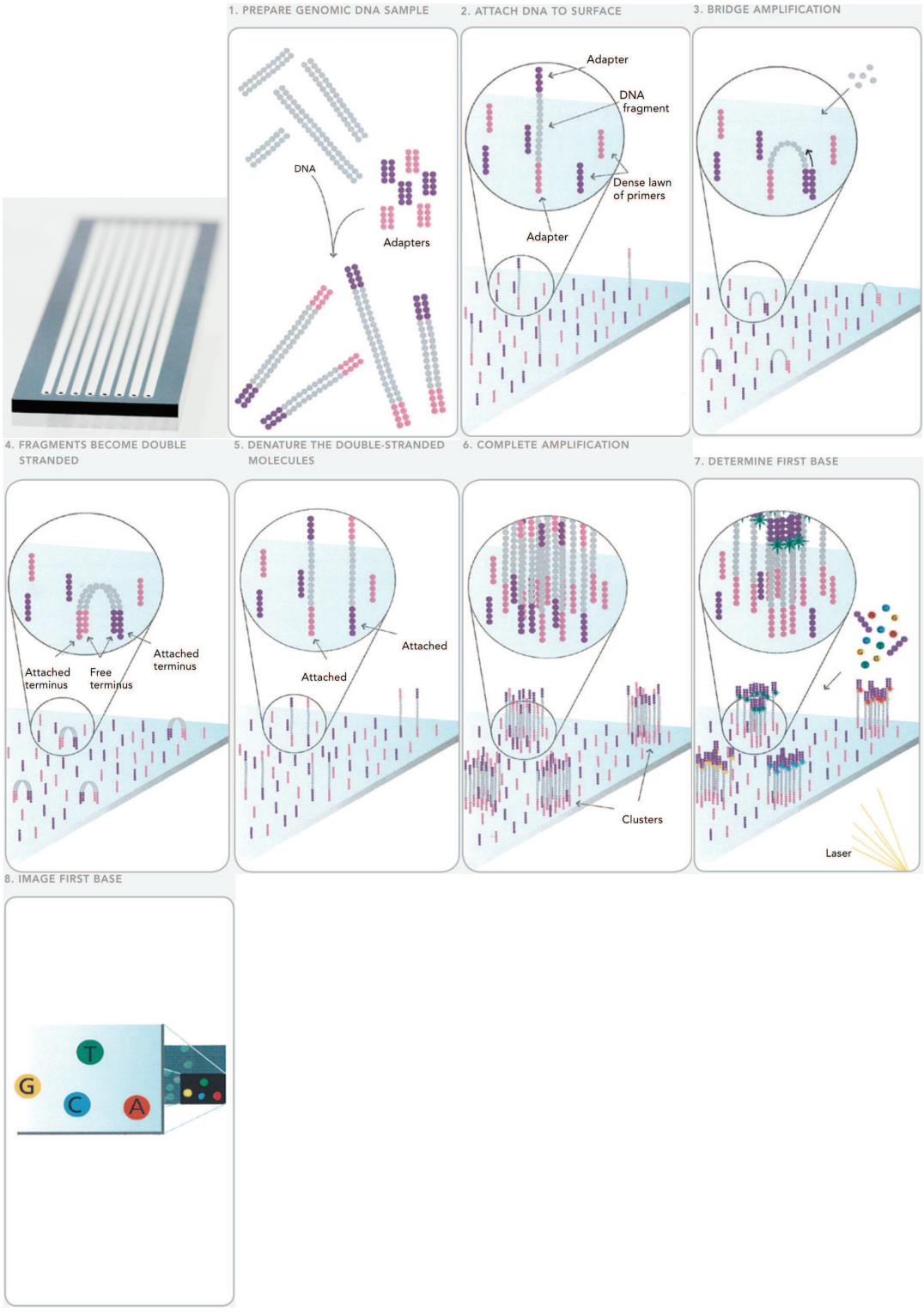


Fig. 2.1 Various steps involved in Illumina sequencing.
Up to eight samples can be loaded onto the flow cell of Genome Analyzer for simultaneous analysis. Various steps from sample preparation to images of bases are illustrated above.
Source (http://www.illumina.com/downloads/GenomeAnalyzerBrochure.pdf)

With the capacity of pooling many samples in one experiment and generating huge amount of data (Valentini *et al.* 2009a), in addition to short sequencing read lengths, these platforms are even ideal for ecological studies, which mostly relay on short and degraded DNA recovered from non-invasive sampling. There has been a growing trend by ecologists to exploit NGS platforms in diet analyses of various species (e.g. Deagle *et al.* 2009, 2010; Pegard *et al*. 2009; Soininen *et al.* 2009; Valentini *et al.* 2009a; Raye *et al.* 2011; Zeale *et al*. 2011; Shehzad *et al.* in press). Next Generation Sequencing revolutionized the analysis of diets based on feces, because the various amplicons of a PCR product can be characterized simultaneously without cloning. Moreover, short DNA fragments either from stomach contents or feces can be easily amplified (Valentini *et al.* 2009b). The increasing availability of NGS platforms at competitive price makes possible large-scale diet studies at low costs (Deagle *et al*. 2009; Pegard *et al.* 2009; Valentini *et al*. 2009a). Recent studies using NGS have marked an epoch in the realm of herbivores diet. Valentini *et al.* (2009a) have presented a universal approach by combining DNA barcoding and pyrosequencing for the diet analysis of herbivores, using a part (*g* and *h*) of the P6 loop of the choloroplast *trn*L (UAA) intron (Taberlet *et al.* 2007) as a barcode. The authors have exploited the full range of plant taxa of various herbivore animals. Comparison with traditional methods with cutting edge techniques evaluated NGS diet assessment as a better choice (Soininen *et al*. 2009). Pegard *et al.* (2009) clearly demonstrated that the NGS approach is much more efficient than cloning to study the diet of herbivores.

2.3 Carnivores diet analyses

Several works have been conducted on the diet of invertebrate preadators, to find the trophic interactions between invertebrate predators and their prey (Symondson 2002; Read *et al*. 2006; King *et al*. 2010). However, our intensions mainly focus here to the diet of large carnivores.

Knowing about the prey of large carnivores, which are highly endangered and difficult to observe due to their cryptic nature and remote habitat is quite challenging. These carnivores are keystone species of an ecosystem and have large effects on its functioning despite relatively low biomass (Mills *et al.* 1993; Power *et al.* 1996). They maintain the biodiversity through resource facilitation, trophic cascade among

different prey species, dependence on ecosystem productivity and extend links to many ecosystem components (Sergio *et al*. 2008).

2.3.1 Role of diet studies in ecology

Predation is a central interspecific relationship that may be studied to understand ecosystem functioning and evolution. It can be studied by multidisciplinary approaches involving ecological, evolutionary and behavioral sciences. Diet studies help evaluating the resources utilized within an ecosystem (Mills 1992) by characterizing prey selection with regards to prey availability. Information about the diet of individuals within an ecosystem is a good indicator of ecosystem stability. For example, it can be used when studying the introduction of an invasive species or the removal of an existing native species, by measuring the influence of these species on the food web. For example, in large parts of western Alaska the sea otter (*Enhydra lutris*) population was largely affected by the arrival of whales (*Orcinus orca*). By predating extensively on sea otter population, whales greatly reduced their number. This resulted in an increase of invertebrate grazers, which in turn reduced the biomass of kelp (Estes *et al.* 1998). Similarly removal of a native predator can also affect the trophic cascades within an ecosystem. Crooks & Soulé (1999) reported that absence of coyotes (*Canis latrans*) in California increased the meso-predators population like opossums, foxes and house cat in the area, which were previously fed by the coyote. These meso-predators ultimately increased the predation pressure on scrub-breeding song birds and native rodents.

Another aspect of analyzing diet relates with the determination of the health status associated with mal or insufficient nutrition. Such studies have been conducted mainly on captive animals in zoos (e.g. Nijboer & Dierenfeld 1996), but it would be informative to collect the same kind of data in the wild. Transfer of pathogens, for instance, bacteria, viruses and parasites is quite easy when a predator feeds on disease-affected prey. In a recent incidence, on 26 February 2011, a snow leopard adult female was captured in Manthal (Sakardu, Gilgit Baltistan province, Pakistan). The animal died soon after capture and was in devastating condition due to the probable predation of a sheep heavily infested by mange mites, an ecto parasite (http://www.snowleopardnetwork.org/blog/?p=353).

Accurate information about diet is also essential to address human-wildlife conflicts as well as predation on vulnerable prey species. Human-wildlife conflicts are common where large carnivores and sympatric human population reliant upon the same resources such as livestock (e.g. Bauer & Kari 2001; Bagchi & Mishra 2006; Sangay & Vernes 2008; Dar *et al*. 2009; Inskip & Zimmermann 2009). Sometimes people overemphasize predation losses to get more financial compensation (Orga 2008). Also, precise information about the diet can help identifying prey with a vulnerable or endangered status (see discussion of articles I and III). Ultimately meticulous prey information can than be helpful to build balanced conservation strategies taking into account people grievances and vulnerability of predator and prey species.

2.3.2 Traditional approaches for carnivores diet

2.3.2.1 Direct or indirect monitoring of predation

Various approaches have been used to study the animal's diet depending upon their potentials and limitations (Shrestha & Wegge 2006). The simplest approach is the direct monitoring of an animal predation in the field (Scheiffarth 2001). This method whenever, possible is more accurate and provides a direct information about the diet of an animal. Moreover, information about the age and sex of the prey species consumed can be assessed. However, this observation is prone to several potential limitations like the presence of an observer, may alter the predation behavior (Gordon 1994). It is laborious because only a limited number of individuals can be monitored at a time. Moreover, this method is impracticable under complex situations, e.g. when an animal feeds on a group of different prey species at a time and leaves no remain for identification (Shehzad *et al.* in revision). Also, it is not easy to observe the diet of nocturnal animals and requires special night vision apparatus (e.g. Allison & Destefano 2006). Moreover, it is nearly impossible to directly note the prey intake when animal is highly elusive and inhabiting remote terrains (e.g. snow leopard; McCarthy *et al.* 2008; Ale & Brown 2009).

Using video monitoring to identify both predator and prey species provides the accurate information on diet without disturbing the animal (Varley *et al.* 1994). This approach has been used to study the diet both in controlled and field conditions

(Merfield *et al.* 2004). But video techniques could be insufficient in field conditions because of the low rate of predators and prey entering into the field of view.

Similarly, radio telemetry is now being used to study the predation of species that are difficult to observe directly. Radio tagged animals are followed and prey estimation is done on the basis of prey remains found. The major advantage of this approach is that information about the age and sex of the prey item consumed is possible. This approach has been used to study the predation behavior of some elusive species like snow leopard (Grongberg 2011), wolf (Knopff *et al*. 2009) and cougar (Anderson & Linzey 2003; Sand *et al*. 2005). Current GPS collaring studies of snow leopards in Mongolia (Panthera/Snow Leopard Trust unpublished data) have been successful in identifying kill sites and prey remains in over 240 instances between 2008 and 2011. That study is unique in its success and is not easily replicated broadly across the range and diet information from such studies remains site specific. Furthermore, following the remains of killed prey in remote areas is extremely difficult (Jackson 1996) and cannot be evident when predator consumes entire prey, leaving no remain. Moreover, radio collaring requires the capture and immobilization of the animal, which is an invasive approach that cannot be encouraged in all situations.

The inference from field surveys, questionnaires and interviews with local community can give information about predation for secretive and highly endangered carnivores (Mishra 1997; Namgail *et al.* 2007; Dar *et al.* 2009). However, this information may be biased towards the public opinion if lacking scientific validation. Most of the time, such surveys are limited only to livestock predation and usually do not provide adequate information about wild prey (see article III).

2.3.2.2 Different methods to study the diet using invasive sampling

Analysis of stomach contents obtained by killing the animal, or after immobilizing and stomach flushing (Wilson 1984) has been widely used. Some studies have also focused the diet of large endangered carnivores by examining stomach contents from dead carcasses (e.g. otter from England; Britton *et al.* 2006 and red fox from Italy; Balestrieri *et al.* 2011). Given that soft-bodied prey items are mostly underestimated in fecal contents, the real assessment of prey items consumed irrespective to their digestibility could be the obvious advantage of this approach (Hyslop 1980). Then stomach contents have been widely used for prey identification

either on micro-histological examination (Soininen *et al.* 2009; Prado *et al*. 2010; Balestrieri *et al.* 2011) or on gross morphological basis.

Besides the visual assessment of prey remains, protein electrophoresis (Greenstone 1999; Symondson 2002) coupled with staining for enzyme activity (e.g. isomerase and estrases; Traugott 2003) has been extensively used to study the diet of invertebrates (Symondson 2002). The extracts from homogenized predators (or either their gut contents) has been analysed on polyacrylamide gels followed by staining for their enzymatic activities. Bands are then compared with those of target prey. The major limitation of this method is its low resolution due to insufficient species-specific diagnostic bands. Furthermore, it is difficult to interpret complex band patterns obtained when a gut contains the remains of several prey species (Walrant & Loreau 1995).

Stomach/gut contents were thoroughly examined using different immunoassays developed for mainly invertebrate carnivore species (see Symondson *et al*. 1999, 2002; Symondson 2002; Fournier *et al*. 2008). Such assays are based on the principle of specific antigen-antibody coupling interactions. The product followed by such interactions is determined by an enzyme-substrate indicator system. Immunoassays can detect the prey protein, which serves as antigens, in the stomach contents of the carnivore. For example, we can use monoclonal antibodies binding with high specificity to an epitope (binding site of the antigen), or polyclonal antibodies binding to a whole antigen. Different immunoassays from the less sensitive protein precipitation test (Berth & Delanghe 2004) and western blotting (Hoyt *et al*. 2000) to the more sensitive enzyme-linked immunosorbent assays (ELISA) have also been used to study the diet of various invertebrates. In western-blot technique, the antigen can be probed with the specific antibodies thus the reaction can be revealed with 5-Bromo-4chloro-3-indolyl phosphate (BCIP) on a nitrocellulose membrane. In ELISA, a positive reaction (specific antigen-antibody interaction) is revealed by the color produced by a substrate/enzyme reaction. ELISA based analysis is quite sensitive and may detect the presence of a prey protein even in minute amount, this method is also quantitative (Naranjo *et al.* 2001). The monoclonal antibody assays were quite useful to detect specific target prey species for which they were developed (Pompanon *et al*. in press). But the main limitation of these assays is the high rate of false positives, due to shared antigenic determinants between various putative prey species (Feller *et*

al. 1985). Moreover, developing monoclonal antibodies is time-consuming and requires laborious hybridoma screening protocols (Sheppard & Harwood 2005).

However, the analysis of stomach/gut contents provides an invasive approach that requires direct handling of animals. So it may be suitable for small or experimental animals but cannot be encouraged for large and endangered animals.

2.3.2.3 Different methods of diet analyses using non-invasive sampling

Feces may represent readily available and easily collected source of information (Putman 1984) without disturbing or touching the animal. This is the method of choice while dealing with endangered and elusive animals (Kohn & Wayne 1997). The main limitation of this sampling technique is that in determining the diet, it represents the short-term diet profile (Deagle 2006). Carnivore feces contain diagnostic and undigested parts, such as bony parts, shells or hairs that help recognizing the prey consumed (Tsukahara 1993; Bowen 2000; Pickering 2001; Tollit *et al.* 2003). However, recovery of hard parts and undigested remains is not always easy, large bones and teeth are generally fragmented and not evident to identify (Oli 1993), providing limited information. The analysis of undigested prey hairs in predator feces through the hair mounting technique has been most commonly used to study the diet of elusive and endangered carnivores (Chundawat & Rawat 1990; Lhagvasuren & Munkhtsog 2000; Bagchi & Mishra 2006; Anwar *et al*. 2011; Bianchi *et al*. 2011). Hair slides are compared with reference specimens on the basis of histological examination (e.g. Oli 1993; Mukherjee *et al*. 1994). This method is laborious and time consuming, in terms that scientists may need to prepare a considerable number of slides from each specimen in order to obtain reliable information (Oli 1993). Also, hair structures may vary according to their place on the same animal fur (Oli 1993; Moyo 2005), or can show similar characteristics for several related species (Oli 1993; Hall-Aspland & Roger 2007). Finally, the lack of reference specimens may be the main barrier in accurate diagnosis.

These limitations have been circumvented by the development of techniques that do not rely on the recovery of digestion-resistant parts. The thin layer chromatography (TLC) of fecal bile acids has been used to some extent (Quinn & Jackman 1994; Fernandez *et al.* 1997). Iverson *et al*. (2004) described that fatty acids that are released from ingested lipid molecules during digestion are not degraded. Since a limited

number of fatty acids can be bio-synthesized by animals, it is possible to distinguish between dietary and non-dietary parts. But the major limitation of this technique is that the fatty acids signatures of prey species may very close to that of the predator and making identification difficult (Piche *et al*. 2010).

The characterization of naturally occurring stable isotope ratios is another powerful tool to describe dietary patterns in terrestrial and marine ecosystems (Rundel *et al.* 1989; Lecomte *et al.* 2011). The ratios of stable isotopes of carbon and nitrogen in the predator tissues and feces depend on its diet (DeNiro & Epstein 1978, 1981; Tieszen & Boutton 1989; Darimont & Reimchen 2002). It is useful where two isotopically distinct dietary sources are available to consumers (Robbins *et al.* 2005). Stable carbon isotopes can distinguish between marine and terrestrial dietary protein (DeNiro & Epstein 1978, 1981), and between herbivores and carnivores protein (Roth & Hobson 2000). Stable nitrogen isotopes reflect the trophic position of an organism within a food web (Ambrose & DeNiro 1986). Isotopic dietary estimates are based on the assimilation and not ingestion of food (Hobson & Clark 1992). The main advantage of this method is that diet can also be inferred from hairs and bones giving long-term information from such samples (Roth & Hobson 2000). This approach can also be used to infer the diet of fossil remains (Feranec & MacFadden 2000; Pérez-Claros & Palmqvist 2008; Garcia *et al.* 2009). The variation of isotopic values of carbon and nitrogen measured in bones or teeth give indications on carnivorous and herbivorous diets. But identification at the species level is not possible through this method.

Near Infrared Reflectance Spectroscopy (NIRS) is another method increasingly used to infer some chemical constituents present in the diet (Park *et al.* 1998). Near infrared spectra depend upon the number of different chemical bonds like O-H, C-H and N-H, which are present in the diet (Dixon 2009). The spectral features, after calibration, are used to predict the composition from unknown samples (Osborne *et al*. 1993; Tran *et al*. 2010). This method is both time and cost effective and reduces the bias between estimated and reference value, as hundreds of samples can be run per day without requiring reagents once the calibration has been achieved (Foley *et al*. 1998). Most commonly NIRS is used in studying foraging strategies in ecology by estimating the nutritional components, including total nitrogen, moisture, fiber, and starch etc. This method has several limitations. Variation and non-homogeneity in particle size

can lead to biased analysis: A larger particle size results in an increased path length for the incident light and thus an increased reflectance. The calibration is a crucial and challenging step that requires a lot of skills and time, and developing a functional calibration for a field study must use samples of those species the predator is considered to be eating (Kaneko & Lawler 2006).

2.3.3 PCR-based/ DNA-based methods for carnivore diet analyses

Molecular analysis provides a better approach to study the diet compared with traditional methods, as it does not rely upon morphological characteristics that may damaged during capturing, ingestion and digestion (Jarman *et al*. 2002; King *et al.* 2008). This approach also provides an improved resolution to identify the prey items, when compared with traditional methods (Soininen *et al*. 2009; Braley *et al*. 2010; Zeale *et al*. 2011). The molecular analysis of diet exploits a wide range of sampling sources, from stomach or gut contents (Deagle *et al*. 2005a; Troedsson *et al*. 2009; King *et al.* 2010; Marshall *et al*. 2010), to regurgitates (Taberlet & Fumagalli 1996) and is even efficient for studying degraded feces (Pegard *et al*. 2009; Valentini *et al*. 2009a; Kowalczyk *et al.* 2011; Raye *et al.* 2011; Shehzad *et al.* in press). It has been extensively used to study a broad range of predators, including marine mammals (Deagle *et al*. 2005b; Dunshea *et al*. 2008; Toillit *et al*. 2009), marine invertebrates (Blankenship & Yayanos 2005; Braley *et al*. 2010), seabirds (Deagle *et al*. 2007a, 2010), terrestrial invertebrates (Garros *et al*. 2008), terrestrial birds (Valentini *et al.* 2009a; Oehm *et al*. 2011), small terrestrial vertebrates (Soininen *et al*. 2009), flying mammals (bats; Bohmann *et al*. 2011; Zeale *et al*. 2011), and also large terrestrial herbivorous and carnivorous vertebrates (Pegard *et al*. 2009; Valentini *et al*. 2009a; Kim *et al*. 2011; Raye *et al*. 2011; Shehzad *et al*. in press).

DNA based approaches are particularly suitable for studying the feeding ecology of elusive and secretive animals, by using their feces as a source of DNA (Deagle *et al*. 2007b; Fernandes *et al*. 2008; King *et al*. 2008; Corse *et al*. 2010). The ideal situation consists in the use of universal primers, as this requires no *a priori* knowledge of the diet to amplify prey DNA. If this approach has been successfully implemented for herbivore (Valentini *et al.* 2009a), the analysis of carnivore diet represents a real challenge, as predator DNA can be simultaneously amplified with prey DNA (Jarman *et al*. 2006; Deagle *et al*. 2007a). Furthermore, prey fragments might be rare in the

DNA extract, and consequently they have the tendency of being missed during the early stages of PCR, resulting in a PCR product containing almost exclusively the dominant sequence of predators (Jarman *et al.* 2004, 2006; Green & Minz 2005). Various methods have been proposed to avoid the amplification of dominant fragments of DNA predator and let the lesser fragments of prey be amplified.

2.3.3.1 Amplifying prey from a specific group or species

Species-specific or group-specific primers have been specially designed to avoid priming on predator DNA and to specifically amplify the target prey species (Vestheim *et al*. 2005; Deagle *et al*. 2006; Juen & Traugott 2007; King *et al.* 2010; Zeale *et al*. 2011). This approach can only be implemented with *a prioiri* knowledge about the potential prey and hence cannot be used in a blind way. Moreover, this is not a suitable method if the potential range of prey species is large (Vestheim & Jarman 2008). Group-specific primers targeting broader groups of prey have been designed in several studies (Walter *et al.* 2000; Jarman *et al*. 2004; 2006; Kuusk & Agusti 2008; Metzler *et al.* 2009). However, designing group-specific primers is not always straightforward when the conserved sequences that can act as priming sites for prey are too close to the homologous sequences of the predator.

The primers targeting a broad range of prey taxa but amplifying a barcode with low taxonomic resolution can be best employed in combination with primers targeting less taxa but leading to higher resolution (Pompanon *et al*. in press). In such cases, the specific primers focus on a particular group with a lower taxonomic level (i.e., genus or species). One such example is the study of Australian fur seal diet (Deagle *et al*. 2009). The authors used a primer set that amplified a wide range of prey sequences at high taxonomic levels. Then more specific primer sets were used to identify particular prey groups with a higher resolution. Another example is the snow leopard diet analysis that used universal primers for vertebrates (*12SV5*; see article III). These universal primers amplify a barcode with high taxonomic resolution (it unambiguously identifies 77 % species and 89 % genus for 103 EMBL releases), but that cannot differentiate domestic sheep (*Ovis aries*) from argali (*Ovis ammon*). We designed a specific primer pair (*Ovis*) using *cytochrome b* gene of mitochondrial DNA that could clearly differentiate between these two prey down to species level.

2.3.3.2 Amplifying all prey with universal primers

As discussed in the previous section, the overwhelming dominance of predator DNA within an environmental sample impedes the amplification of all prey DNA when a PCR is done with universal primers. However, various approaches have been proposed either by using restriction enzymes or by modifying the oligonucleotides, to limit the amplification of predator DNA.

A restriction enzyme chosen to digest the target fragment of the predator DNA and leaving the DNA of all prey intact may serve the purpose. After such digestion the predator DNA is no longer available for PCR amplification and all prey target fragments (prey) can be amplified with universal primers (Blankenship & Yayanos 2005). This method has been successfully applied to reveal the prey diversity in various studies (e.g. Blankenship & Yayanos 2005; Suzuki *et al*. 2006, 2008; Dunshea 2009). However, this method has several technical constraints. First, this methodology will only be effective when both predator and prey sequences are sufficiently different from each other; otherwise, the restriction enzymes, which cut the predator sequences, can also digest the prey sequences (Green & Minz 2005). Second, the DNA must be under its double stranded form for restriction enzyme to digest it, and single stranded predator fragments can be avoided by restriction enzymes and subsequently amplified. Third, this approach is not effective once predator sequences have already been amplified because the DNA of several prey may be already missing and could not be retrieved even with another subsequent amplification. This problem was addressed by Dunshea (2009) who tested the effect a digestion enzyme treatment before and after PCR to avoid overcoming predator amplicons. Fourth, this approach still requires *a priori* knowledge of all possible prey sequences to be sure that restriction enzyme will not digest any prey DNA.

The PCR clamping approach uses a modified oligonucleotide (a Peptide Nucleic Acid, PNA) to suppress the dominant sequences present in a mixture. A PNA is a synthetic analogue of DNA (Deoxyribonucleic acid) generating PNA-DNA duplexes that are thermodynamically more stable than DNA duplexes (Fig. 2.2a) (Ørum *et al*. 1993; Nielsen & Egholm 1999; Sawata *et al*. 1999). Thus, a PNA analogue specific of a sequence preferentially anneals with it, making this complex inert for elongation by a polymerase, as a PNA cannot function as a primer (Ørum 2000). Consequently, the other sequences present in a template, even rare, can be amplified with universal

primers (Chow *et al*. 2011). This tactic has been widely used in various studies to block dominant sequences (e.g. Demers *et al*. 1995; Hyrup & Nielsen 1996; Knudsen & Nielsen 1997; Macadangdang *et al*. 2011) even for diet assessment (Chow *et al*., 2011). However, this approach is not cost effective. It is six times more expensive than a blocking oligonucleotide used for the same purpose. Another potential problem could be the non-specific clamping of PNA (Chow *et al*. 2011).

Another kind of modified oligonucleotide (Locked Nucleic Acid, LNA) has been used to tightly lock a specific dominant sequence and restrain its amplification. LNA are synthetic analogues in which the ribose ring is "locked" by a methylene bridge between the 2-oxygen and the 4-carbon (Fig. 2.2b), thus reducing the conformational flexibility of the ribose and increasing binding affinity within the phosphate (Braasch & Corey 2001). Several studies have demonstrated that LNA has strong affinity and sensitivity for its complementary strand of DNA (Dominguez & Kolodney 2005; Najafi-Shoushtari *et al*. 2010). Addition of each locked nucleobase in an oligonucleotide can increase its melting temperature (Tm) by 3-6°C. Hu *et al*. (2009) used LNA protocol to detect point mutation in cancerous versus healthy cells. Addition of a locked nucleobase is an expensive procedure thus the method is not cost effective. Furthermore, Tm of the reaction may increase with addition of each locked nucleobase.

Dual Priming Oligo (DPO) is another technique based on a modified oligonucleotide system, which can be applied to study rare sequences within a template. The DPO is a double priming system consisting of a long primer at the 5' end (stabilizer), a short primer at the 3' end (determiner), and a polydeoxyinosine linker (Fig. 2.2c) (Chun *et al.* 2007; Yoo *et al.* 2007). The two primers have distinct priming temperatures while the linker makes a bubble like structure and is not involved in the priming, which does not increase the Tm of the oligonucleotide (Chun *et al.* 2007). Due to its high reliability and sensitivity, the DPO platform has been widely used in the detection of pathogens and rare mutation in various clinical samples (Woo *et al*. 2008; Horii *et al*. 2009; Kwak *et al*. 2010). Vestheim & Jarman (2008) applied this DPO system to study the diet of Antarctic krill. In order to block amplification of krill DNA, the authors designed a DPO specific of the krill target sequence. During a PCR reaction with universal primers this system has significantly prevented the amplification of predator's DNA allowing that of prey DNA.

The simplest and cheapest of all modified oligonucleotides that can be used to stop the amplification of a given sequence includes a 3-carbon spacer (C3-spacer) (Fig. 2.2d). This blocking oligonucleotide will hybridize with the complementary target DNA, making it inaccessible for the polymerase. The rare sequences present in a template can then be amplified with universal primers (Vestheim & Jarman 2008). This method has been effectively used in many scientific domains including clinical chemistry (Kageyama *et al.* 2008; Wang *et al.* 2008; Li *et al.* 2007, 2009), environmental microbiology (Liles *et al.* 2003) and even in archeological sciences (Gigli *et al*. 2009). However, its use in dietary analysis opens broad perspectives. Vestheim and Jarman (2008) first used a blocking oligonucleotide to assess the diet of Antarctic krill. More recently, Deagle *et al.* (2009, 2010) investigated the diet of Australian fur seals (*Arctocephalus pusillus*) and penguins (*Eudyptula minor*) by combining this blocking oligonucleotide approach with the use of the 454 GS-FLX pyrosequencing technology for subsequent characterization of PCR products.

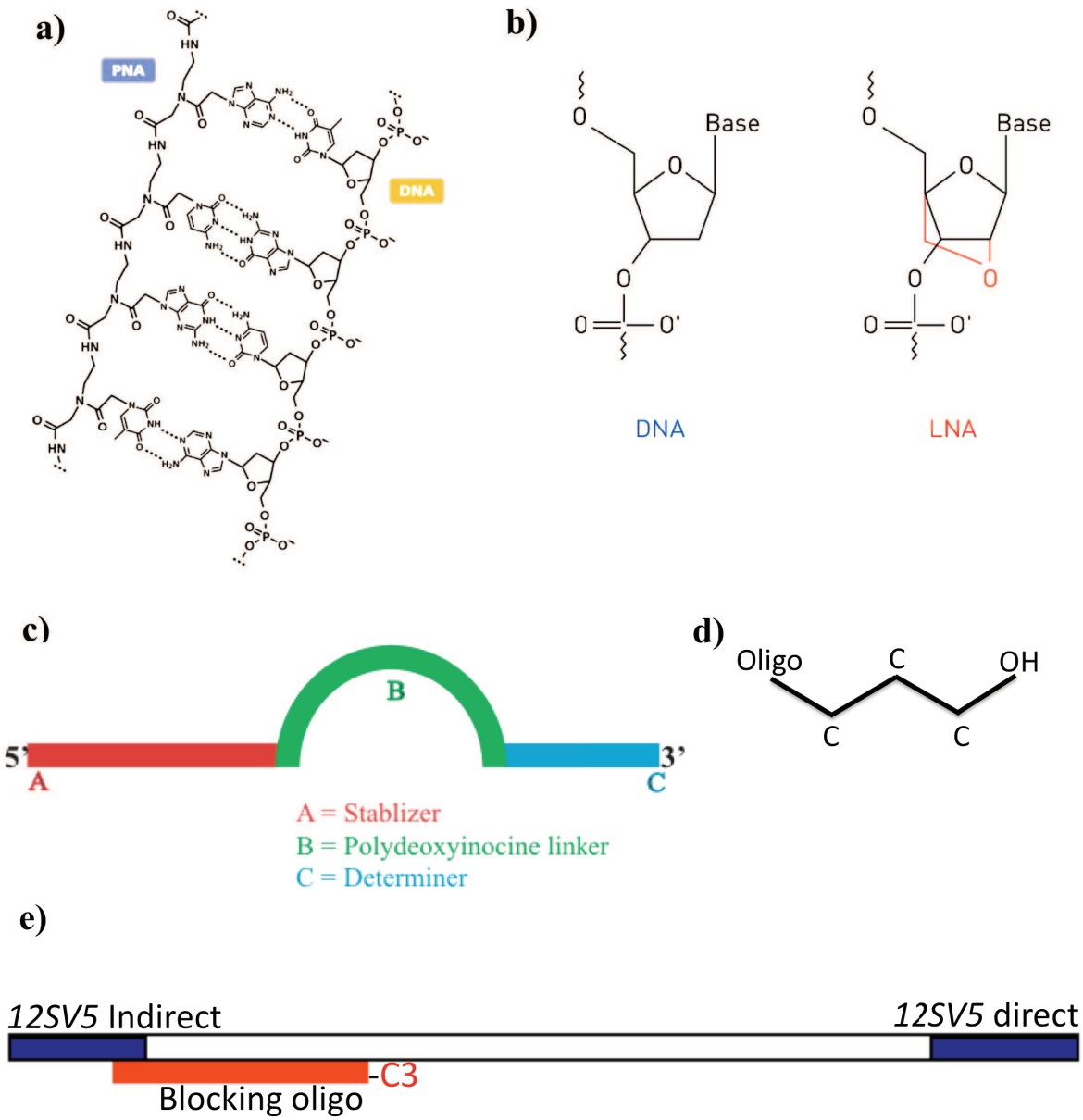


Fig. 2.2 Various modified oligonucleotides

a) Peptide nucleic acid, the normal phosphate backbone of a DNA is replaced by a polyamide backbone **b)** Locked nucleic acid, the ribose ring is "locked" by a methylene bridge between 2-oxygen and 4-carbon **c)** Dual priming oligonucleotide, having a long primers (stablizer), a short primer (determiner) and a polydeoxyinosine linker **d)** Carbon3 spacer (C3-spacer) is attached at the 3' end of an oligonucleotide to block complementry sequences **e)** C3 spacer (designed for this study)

2.4 Potentials of metabarcoding diet assessement to adress ecological and social issues

DNA metabarcoding has many advantages over the other methods previously discussed in this chapter, to resolve several ecological and social issues. As discussed in the section 2.3.1, DNA metabarcoding helps to answer the ecological questions related to the introduction of an alien, or removal of a native species within an ecosystem. The impact on food web can be measured by exploiting environmental samples, and such an approach is ideal for wild and elusive species without observing or capturing them.

Similarly, metabarcode approaches have potentials to assess the real extent of human-carnivore conflicts. Such conflicts arise, when human and carnivores live in sympatry and share the same resources (Inskip & Zimmermann 2009). The causes and nature of such conflicts have been discussed in detail in articles (II and III). The metabarcode approaches are helpful to accurately analyse the diet and hence contribute to resolve these conflicts. For examples, when a predator is held responsible relying on livestock, then measures to mitigate such conflicts should be adopted, which would address both community grievances and conservation of a carnivore in the area (see discussion of article II). Conversely, the accurate diet assessment helps to reduce the unnecessary stress to a carnivore, from the local community when it is just alleged to predate on livestock, but in real sustaining on wild prey (article III).

2.5 Objective of this study

This thesis has following two main objectives

1- To develop a universal approach for carnivores diet analysis, without any *a priori* information about all the vertebrate prey taxa (**article I**). We combined metabarcoding approach with next generation sequencing to achieve this objective.

Using feces as a source of DNA we have designed universal primers for vertebrates *12SV5* (Riaz *et al*. 2011). To limit the amplification of predator sequences which normally are in excess or sometimes predominated the PCR products, predator specific blocking oligonucleotides were designed. These blocking oligonucleotides were modified at 3' end with a C3 spacer (see Fig. 2.2d & Fig. 2.2e). We validated the effect of the blocking oligonucleotide by performing PCRs with the *12SV5F/12SV5R* primer pair in absence or presence of a predator specific blocking oligonucleotide (20 times concentrated than the PCR primers). The amplicons were subsequently sequenced using a NGS platform to reveal the diet.

2- Apply this approach to other carnivores diet: common leopard (*Panthera pardus*) population of Pakistan **(article II)**, snow leopard (*Panthera uncia*) population from Mongolia **(article III)**.

We studied three endangered wild felids (leopard cat, common leopard and snow leopard) from Asia, for which different scientific questions related to conservation issues were addressed: (i) the intraspecific variability of the diet, between two Pakistani populations of leopard cat, (ii) Human-leopard conflicts, when predator relies mainly on domestic animals for its survival, (iii) the assessment of snow leopard diet, when it is mostly alleged for predation on livestock prey.

MOLECULAR ECOLOGY

Molecular Ecology (2012) doi: 10.1111/j.1365-294X.2011.05424.x

Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan

WASIM SHEHZAD,* TIAYYBA RIAZ,* MUHAMMAD A. NAWAZ,*† CHRISTIAN MIQUEL,* CAROLE POILLOT,* SAFDAR A. SHAH,‡ FRANÇOIS POMPANON,* ERIC COISSAC* and PIERRE TABERLET*

**Laboratoire d'Ecologie Alpine, CNRS-UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 9, France, †Snow Leopard Trust (Pakistan Program), 17-Service Road North, I-8/3, Islamabad, Pakistan, ‡Wildlife Department, Khyber Pakhtunkhwa, Pakistan*

Abstract

Diet analysis is a prerequisite to fully understand the biology of a species and the functioning of ecosystems. For carnivores, traditional diet analyses mostly rely upon the morphological identification of undigested remains in the faeces. Here, we developed a methodology for carnivore diet analyses based on the next-generation sequencing. We applied this approach to the analysis of the vertebrate component of leopard cat diet in two ecologically distinct regions in northern Pakistan. Despite being a relatively common species with a wide distribution in Asia, little is known about this elusive predator. We analysed a total of 38 leopard cat faeces. After a classical DNA extraction, the DNA extracts were amplified using primers for vertebrates targeting about 100 bp of the mitochondrial 12S rRNA gene, with and without a blocking oligonucleotide specific to the predator sequence. The amplification products were then sequenced on a next-generation sequencer. We identified a total of 18 prey taxa, including eight mammals, eight birds, one amphibian and one fish. In general, our results confirmed that the leopard cat has a very eclectic diet and feeds mainly on rodents and particularly on the Muridae family. The DNA-based approach we propose here represents a valuable complement to current conventional methods. It can be applied to other carnivore species with only a slight adjustment relating to the design of the blocking oligonucleotide. It is robust and simple to implement and allows the possibility of very large-scale analyses.

Keywords: blocking oligonucleotide, DNA metabarcoding, mitochondrial DNA, ribosomal DNA, species identification

Received 8 June 2011; revision received 16 November 2011; accepted 24 November 2011

Introduction

The nature of trophic interactions is a fundamental question in ecology and has commanded the attention of biologists for decades. Dietary behavioural studies provide key data for understanding animal ecology, evolution and conservation (Symondson 2002; Krahn *et al.* 2007). Wild felids are among the keystone predators and have significant effects on ecosystem functioning, despite their relatively low biomass (Mills *et al.* 1993; Power *et al.* 1996). The modal mass concept (Macdonald *et al.* 2010) proposes that each felid species focuses on large-as-possible prey to maximize their intake relative to their energy expenditure for each catch, provided that such prey can be safely killed.

Owing to their elusive behaviour, scientific knowledge of South Asian wild cats is limited (Nowell & Jackson 1996). The leopard cat (*Prionailurus bengalensis*) is a small felid (weight 1.7–7.1 kg; Sunquist & Sunquist 2009), with a wide range in Asia (8.66×10^6 km^2; Nowell & Jackson 1996). Beginning in Pakistan and

Correspondence: Pierre Taberlet, Fax: +33(0)4 76 51 42 79; E-mail: pierre.taberlet@ujf-grenoble.fr

parts of Afghanistan in the west, the leopard cat occurs throughout Southeast Asia, including the islands of Sumatra, Borneo, and Taiwan. It extends into China, Korea, Japan and the Far East of Russia. (Macdonald *et al.* 2010). The leopard cat's flexible habitat selection and prey choices favour its distribution throughout the range (Watanabe 2009; Mukherjee *et al.* 2010). It is found in very diverse environments, from semideserts to tropical forests, woodlands to pine forests and scrubland to agriculture land (Sunquist & Sunquist 2002). It prefers to live in habitats near sources of water and can be found in the close proximity to human population (Scott *et al.* 2004).

The population status of the leopard cat is not uniform throughout its range. The cat is relatively secure in China (Lau *et al.* 2010) and in India (Nowell & Jackson 1996), endangered in Korea (Rho 2009) and most endangered in Japan (Mitani *et al.* 2009). In Pakistan, this species is categorized by the IUCN as ''data deficient'' as no information exists about the extent of its occurrence, nor its occupancy, population and habitat (Sheikh & Molur 2004). Major threats to the species include hunting, habitat loss and fragmentation because of the human population expansion in addition to competition for prey with other sympatric carnivores (Izawa & Doi 1991). Commercial exploitation for the fur trade is a significant threat throughout its range (Sheikh & Molur 2004); in China, the annual pelt harvest was estimated at to be 400 000 animals in mid-1980s (Nowell & Jackson 1996).

Despite being a relatively common species with a wide distribution, comparatively little information is available about the diet of the leopard cat in general, and no information at all specific to Pakistan, where this predator is rare. Faeces analysis by hair mounting and bone examination is used extensively and can provide information about the diet (e.g. Oli *et al.* 1994; Gaines 2001; Bagchi & Mishra 2006; Lovari *et al.* 2009). Muridae (mainly *Rattus* spp. and *Mus* spp.) seem to represent the main prey items throughout the leopard cat distribution range, supplemented by a wide variety of other prey including small mammals such as shrews and ground squirrels, birds, reptiles, frogs and fish (Tatara & Doi 1994; Grassman *et al.* 2005; Austin *et al.* 2007; Rajaratnam *et al.* 2007; Watanabe 2009; Fernandez & de Guia 2011).

Molecular analysis of faeces (Höss *et al.* 1992; Kohn & Wayne 1997) provides an alternative noninvasive approach to study animal diet, but prey DNA in faeces is often highly degraded, preventing the amplification of long fragments (Zaidi *et al.* 1999; Jarman *et al.* 2002). Until 2009, most of the molecular-based studies to analyse diet were carried out using traditional sequencing approaches (e.g. Deagle *et al.* 2005a, 2007; Bradley *et al.* 2007). These methods require cloning PCR products and subsequent Sanger sequencing of these clones by capillary electrophoresis. However, this approach is both time-consuming and expensive (Pegard *et al.* 2009).

Next-generation sequencing is revolutionizing diet analysis based on faeces (Valentini *et al.* 2009b), because sequence data from very large numbers of individual DNA molecules in a complex mixture can be studied without the need for cloning. Valentini *et al.* (2009a) have presented a universal approach for the diet analysis of herbivores. The methodology consists of extracting DNA from faeces to amplify it using the universal primers *g* and *h,* which amplify the short P6 loop of the chloroplast *trn*L (UAA) intron (Taberlet *et al.* 2007), and in sequencing the PCR products using a next-generation sequencer.

While such an approach has been successfully implemented for herbivores, the analysis of carnivore diet presents a real challenge when using primers for mammals or vertebrates, as predator DNA can be simultaneously amplified with prey DNA (Deagle *et al.* 2005b; Jarman *et al.* 2006). Furthermore, prey fragments might be rare in the DNA extract from faeces, and consequently be prone to being missed during the early stages of PCR, resulting in a PCR product almost exclusively containing the dominant sequences of predators (Jarman *et al.* 2004, 2006; Green & Minz 2005). Various methods have been proposed to avoid amplifying predator DNA. Species-specific or group-specific primers have been specially designed to avoid priming on predator DNA and to specifically amplify the target prey species (Vestheim *et al.* 2005; Deagle *et al.* 2006; King *et al.* 2010). This is not a convenient strategy if the prey are taxonomically diverse, which makes the design of suitable primers difficult (Vestheim & Jarman 2008). Another strategy involves cutting predator sequences with restriction enzymes before and/or during and/or after PCR amplification (Blankenship & Yayanos 2005; Green & Minz 2005; Dunshea 2009). However, these approaches can only be implemented with *a priori* knowledge of the potential prey.

The ideal system for studying carnivore diet using DNA in faeces lies in combining, in the same PCR, primers for vertebrates and a blocking oligonucleotide with a 3-carbon spacer (C3-spacer) on the 3′ end that specifically reduces the amplification of the predator DNA. Such a blocking oligonucleotide must be specifically designed to target predator DNA and thus bind preferentially with predator sequences, limiting their amplification. This concept has been effectively used in the field of clinical chemistry (Kageyama *et al.* 2008; Wang *et al.* 2008; Li *et al.* 2009) and in environmental microbiology (Liles *et al.* 2003). However, the application of blocking oligonucleotide in trophic studies is relatively recent. Vestheim & Jarman (2008) first used a

blocking oligonucleotide to assess the diet of Antarctic krill. More recently, Deagle *et al.* (2009, 2010) investigated the diet of Australian fur seals (*Arctocephalus pusillus*) and penguins (*Eudyptula minor*) by combining a blocking oligonucleotide approach with 454 GS-FLX pyrosequencing technologies.

The main aim of this study was to analyse the leopard cat diet in two distinct environments in Pakistan by developing a method that would give the vertebrate diet profile of a carnivore without any *a priori* information about the prey species. This method is based on the use of recently designed primers for vertebrates (Riaz *et al.* 2011) together with a blocking oligonucleotide specific to the leopard cat and employing a high-throughput next-generation sequencer. However, such an approach cannot detect the cases of infanticide and possible cannibalism that have been documented in Felidae (e.g. Natoli 1990).

Materials and methods

General strategy for diet analysis of the leopard cat

Figure 1 outlines the general strategy we followed for the diet analysis of the leopard cat. After the faeces collection and DNA extraction, the samples were confirmed to be those of leopard cat by using leopard cat–specific primers. Selected samples were amplified in two series of experiments, one with primers for vertebrates and the other with the same primers plus a blocking oligonucleotide specific to the leopard cat. These PCR products were subsequently sequenced using the Illumina sequencing platform GA IIx. The amplified sequences of prey taxa were identified by comparison with reference databases (GenBank/EMBL/DDBJ), taking into account prey availability according to their geographic distributions.

Sample collection and preservation

Putative felid faeces were collected in two areas: Ayubia National Park (ANP) and Chitral Gol National Park (CGNP). Both national parks are located in the Khyber Pakhtunkhwa province and represent two extremities of the leopard cat range in Pakistan (Fig. 2). These national parks have disparate environments. The ANP is comprised of moist temperate forests, subalpine meadows and subtropical pine forests. Mean temperatures range between 4.2 °C in January to 26 °C in July. The altitudinal variation ranges from 1050 to 3027 m, and the mean annual rainfall is between 1065 and 1424 mm. It has ~200 species of birds, 31 species of

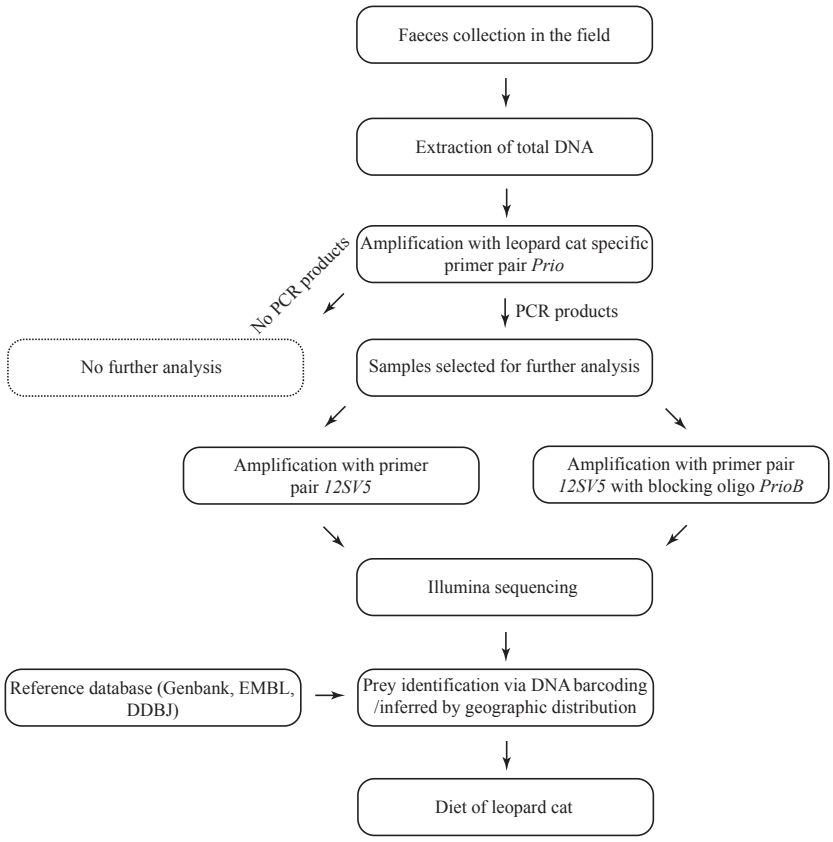


Fig. 1 Flowchart diagram showing the various steps involved in the diet analysis of the leopard cat. The samples in the dotted box were discarded from further experimentation.

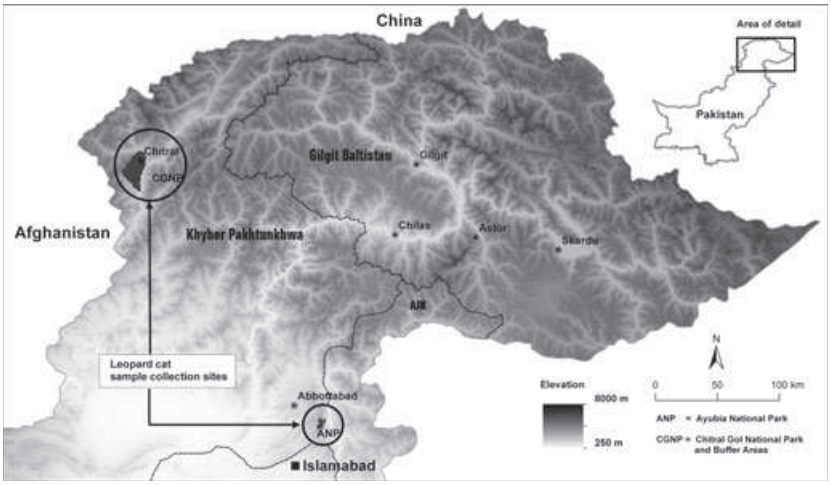


Fig. 2 Sampling locations of leopard cat faeces in northern Pakistan.

mammals, 16 species of reptiles and three species of amphibians (Farooque 2007).

The CGNP generally falls into a subtropical zone with vegetation classified as dry temperate forests. Forests of the park are growing under the extremes of climatic and edaphic factors, and tree canopy is rarely closed. Mean temperature of the valley ranges between 1 °C in January to 24 °C in July, and average annual rainfall varies between 450 and 600 mm. The park supports 29 mammals, 127 birds and nine reptiles (GoN-WFP & IUCN 1996; Mirza 2003).

We collected 114 faecal samples from ANP and 67 from the CGNP. The samples were preserved first in 90% ethanol and then shifted into silica gel for transportation to LECA (Laboratoire d'Ecologie Alpine), Université Joseph Fourier, Grenoble, France, for diet analysis.

DNA extraction

All extractions were performed in a room dedicated to degrade DNA extractions. Total DNA was extracted from about 15 mg of faeces using the DNeasy Blood and Tissue Kit (QIAgen GmbH). Each 15 mg faecal sample was incubated for at least 3 h at 55 °C with a lysis buffer (Tris–HCl 0.1 M, EDTA 0.1 M, NaCl 0.01 M and *N*-lauroyl sarcosine 1% with pH 7.5–8), before following the manufacturer's instructions. The DNA extracts were recovered in a total volume of 250 μL. Blank extractions without samples were systematically performed to monitor possible contaminations.

Selection/designing of primer pairs for the leopard cat diet study

Identification of faecal samples as leopard cat. We used the leopard cat–specific primer pair *PrioF/PrioR*, amplifying a 54-bp fragment (without primers) of the mitochondrial 12S gene (Table 1). The specificity of this primer pair was validated both by empirical experiments (Ficetola *et al.* 2010) and by the program *ecoPCR* (Bellemain *et al.* 2010; Ficetola *et al.* 2010), with parameters to prevent mismatches on the two last nucleotides of each primer, and designed to tolerate a maximum of three mismatches on the remaining part of the primers. The goal of such an experimental validation was to distinguish leopard cat faeces from those from the two other felid species potentially occurring in the study areas, i.e. the common leopard (*Panthera pardus*) in ANP and the snow leopard (*Panthera uncia*) in CGNP. The primary identification of samples was carried out on the basis of the presence of a PCR product of the suit-

Table 1 Sequences of the primer pairs used in the study. The length of amplified fragments (excluding primers) with *Prio* & *12SV5* was 54 and ~100 bp, respectively

Name	Primer sequence (5–3′)	References
PrioF	CCTAAACTTAGATAGTTAATTTT	Ficetola *et al.* (2010)
PrioR	GGATGTAAAGCACCGCC	Ficetola *et al.* (2010)
12SV5F	TAGAACAGGCTCCTCTAG	Riaz *et al.* (2011)
12SV5R	TTAGATACCCCACTATGC	Riaz *et al.* (2011)
PrioB	CTATGCTTAGCCCTAAACTTAGATAGTTAATTTTAACAAAACTATC-C3	This study

able length as revealed by electrophoresis on a 2% agarose gel. The samples successfully amplified using *PrioF/PrioR* were selected for further analyses.

The PCRs were carried out in a total volume of 20 μL with 8 mM Tris–HCl (pH 8.3), 40 mM KCl, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 μM of each primer, BSA (5 μg), 0.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) using 2 μL of DNA extract as a template. The PCR conditions were set as an initial 10-min denaturation step at 95 °C to activate the polymerase, followed by 45 cycles of denaturation at 95 °C for 30 s and annealing at 53 °C for 30 s, without elongation steps as the amplified fragment was very short.

Blocking oligonucleotide specific to leopard cat sequences. The *PrioB* (Table 1) blocking oligonucleotide specific to leopard cat sequences was designed as suggested by Vestheim & Jarman (2008). This blocking oligonucleotide was used to limit the amplification of leopard cat sequences when using the primers targeting all vertebrates. Table 2 presents a sequence alignment of *PrioB* with the main groups of vertebrates. This blocking oligonucleotide might also slightly block the amplification of other felid species, but will not prevent the amplification of other vertebrate groups.

Primer pair for vertebrates. We used the primer pair *12SV5F/12SV5R* designed by the *ecoPrimers* program (Riaz *et al.* 2011). *ecoPrimers* scans whole genomes to find new barcode markers and their associated primers, by optimizing two quality indices measuring the taxonomical coverage and the discrimination power to select the most efficient markers, according to specific experimental constraints such as marker length or targeted taxa. This primer pair for vertebrates represents the best choice found by *ecoPrimers* among short barcodes, as derived from the available vertebrate whole mitochondrial genomes currently available. It amplifies a ~100-bp fragment of the V5 loop of the mitochondrial 12S gene, with the ability to amplify short DNA fragments such as those recovered from faeces, and has a high taxonomic resolution, despite its short size. Using the *ecoPCR* program (Bellemain *et al.* 2010; Ficetola *et al.* 2010), and based on the release 103 of the EMBL database, this fragment unambiguously identifies 77% of the species and 89% of the genera as recorded by this EMBL release (Riaz *et al.* 2011).

DNA amplification for diet analysis

All DNA amplifications were carried out in a final volume of 25 μL, using 2 μL of DNA extract as template. The amplification mixture contained 1 U of AmpliTaq Gold® DNA Polymerase (Applied Biosystems), 10 mM Tris–HCl, 50 mM KCl, 2 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.1 μM of each primer (*12SV5F/12SV5R*) and 5 μg of bovine serum albumin (BSA; Roche Diagnostic). The PCR mixture was denatured at 95 °C for 10 min, followed by 45 cycles of 30 s at 95 °C and 30 s at 60 °C; as the target sequences are ~100 bp long, the elongation step was removed to reduce the +A artefact (Brownstein *et al.* 1996; Magnuson *et al.* 1996) that might decrease the efficiency of the first step of the sequencing process (blunt-end ligation). Using the aforementioned conditions, the DNA extracts were amplified twice, first with *12SV5F/12SV5R* (0.1 μM each) and second with *12SV5F/12SV5R/PrioB* (0.1 μM for *12SV5F* and *12SV5R*, 2 μM for *PrioB*). These primer concentrations have been chosen after a series of test experiments, with various concentrations of *PrioB* (data not shown).

Table 2 Sequence alignment showing the specificity of the *PrioB* blocking oligonucleotide. The first six nucleotides of the *PrioB* blocking oligonucleotide overlap with the *12SV5R* amplification primer. This sequence alignment contains two other Felidae species (*Felis catus* and *Panthera tigris*), another carnivore species from the Ursidae family (*Ursus arctos*), two rodents (*Rattus rattus* and *Microtus kikuchii*), one insectivore (*Crocidura russula*), one bird (*Gallus gallus*), one amphibian (*Rana nigromaculata*) and one fish (*Cyprinus carpio*)

Accession number	Species name	Sequences (5′–3′)
PrioB blocking oligonucleotide		CTATGCTTAGCCCTAAACTTAGATAGTTAATTTTAACAAAACTATC
HM185183	*Prionailurus bengalensis*	.....
NC_001700	*F. catus*	.....CCC.A.....
JF357967	*P. tigris*	.....C.....CCCA.....
NC_003427	*U. arctos*	.....T.....A..A...A..T...AA.CA...TTAT..
NC_012374	*R. rattus*	.....C.TA...A...CA.C..CA....TAT.T
NC_003041	*M. kikuchii*	.....C.TAG..A..AAAAC.A...TA.T.G..
NC_006893	*C. russula*	.....A.A.C.A.C..A.AAC.AG.CTG.TCG
NC_007236	*G. gallus*	.....C.....TC.....CC.CCCA.C.CAC.TGTATC.
NC_002805	*R. nigromaculata*	T.....C.....GT....AATC.ACTCAC.CCAACCA.CGC.AGGG
NC_001606	*C. carpio*	.....C....G.....C...C.TCC.GC.AC..TT.G.TGTC.

The primers for vertebrates, *12SV5F* and *12SV5R*, were modified by the addition of specific tags on the 5′ end to allow the assignment of sequence reads for the relevant sample (Valentini *et al.* 2009a). All of the PCR products were tagged identically on both ends. These tags were composed of CC on the 5′ end followed by seven variable nucleotides that were specific to each sample. The seven variable nucleotides were designed using the *oligoTag* program (http://www.prabi.grenoble.fr/trac/OBITools) to have at least three differences among the tags, to contain no homopolymers longer than two and to avoid a C on the 5′ end so as to allow the detection of a possible deletion within the tag. All of the PCR products from the different samples were first purified using the MinElute PCR purification kit (QIAGEN GmbH), titrated using capillary electrophoresis (QIAxel; QIAgen GmbH) and finally mixed together in equimolar concentration before sequencing.

DNA sequencing

The sequencing was carried out on the Illumina Genome Analyzer IIx (Illumina Inc.), using the Paired-End Cluster Generation Kit V4 and the Sequencing Kit V4 (Illumina Inc.), following the manufacturer's instructions. A total of 108 nucleotides were sequenced on each extremity of the DNA fragments.

Sequence analysis and taxon assignation

The sequence reads were analysed separately with and without the blocking oligonucleotide, using the OBITools (http://www.prabi.grenoble.fr/trac/OBITools). First, the direct and reverse reads corresponding to a single molecule were aligned and merged using the *solexaPairEnd* program, taking into account data quality during the alignment and the consensus computation. Primers and tags were then identified using the *ngsfilter* program. Only sequences with a perfect match on tags and a maximum of two errors on primers were recorded for the subsequent analysis. The amplified regions, excluding primers and tags, were kept for further analysis. Strictly, identical sequences were clustered together using the *obiuniq* program, keeping the information about their distribution among samples. Sequences shorter than 60 bp, or containing ambiguous nucleotides, or with occurrence lower or equal to 100 were excluded using the *obigrep* program. Taxon assignation was achieved using the *ecoTag* program (Pegard *et al.* 2009). *EcoTag* relies on a dynamic programming global alignment algorithm (Needleman & Wunsch 1970) to find highly similar sequences in the reference database. This database was built by extracting the relevant part of the mitochondrial 12S gene from EMBL nucleotide library using the *ecoPCR* program (Bellemain *et al.* 2010; Ficetola *et al.* 2010). A unique taxon was assigned to each unique sequence. This unique taxon corresponds to the last common ancestor node in the NCBI taxonomic tree of all the taxids of the sequences of the reference database that matched against the query sequence. Automatically assigned taxonomic identification was then manually curated to further eliminate those sequences that were the likely result of PCR artefacts (including chimeras, primer dimers or nuclear pseudogenes) or from obvious contaminations. Usually, chimeras can be easily identified by their low identity (<0.9) over the entire query sequence length with any known sequence and by their low frequency when compared with the main prey items. Finally, the prey items were tentatively identified by correlating sequence data with the potential leopard cat vertebrate prey known to be present in the two regions where the faeces were collected, with the constraint that such potential prey must be phylogenetically close to the prey identified in the public database by the *ecoTag* program. The significance of diet differences between ANP and CGNP was assessed by Pearson's chi-squared tests with simulated *P*-values based on 10^6 replicates, using the frequency of occurrence of prey in faeces. Results of such a test have to be analysed carefully because categories used in the contingency table are prey and several prey are detected in each faeces (Wright 2010). This potentially induced a bias if we consider that two prey in the same faeces cannot be considered as independently sampled. If it really exists, the dependency between prey count leads us to overestimate the true number of degrees of freedom. This is a main problem if the test is not rejecting the null hypothesis, but in case of the rejection of this null hypothesis, this places us on the conservative side of the decision.

Rarefaction analysis of prey in faeces originating from ANP and CGNP

We used species rarefaction curve to estimate the total number of prey species likely to be eaten by the leopard cat in the two study areas. The species accumulation, based on the faecal samples, was computed using the analytical formulas of Colwell *et al.* (2004) in EstimateS (Version 8.2, R. K. Colwell, http://purl.oclc.org/estimates).

Results

Of 181 putative felid faeces collected in the field, 38 samples were confirmed to be that of leopard cat with species-specific primers (22 from ANP of 114, and 16

from CGNP of 67). The next-generation sequencing generated about 0.6 and 0.5 million sequences for the samples without and with the blocking oligonucleotide (Table 3), respectively. After applying different filtering programmes, we finally obtained 232 and 141 sequences from the run without and with blocking oligonucleotides, respectively. Sequences within a sample having either a low frequency (e.g. <0.01 when compared with the most frequent sequence) or being very similar to a highly represented sequence were considered to be amplification/sequencing errors and were discarded. All faeces identified as leopard cat with the species-specific primers were confirmed by sequencing. The leopard cat sequence (accession numbers FR873685 and FR873686) was found with a frequency superior to 0.5 in all samples when using only the *12SV5* primer pair (Fig. 3). As in similar experiments (e.g. Deagle *et al.* 2009), we found some human contaminations corresponding to 0.2% and 5.4% of the sequences without and with the blocking oligonucleotide, respectively. A few PCR artefacts with very short sequences were also observed when using the blocking oligonucleotide, but not without blocking.

Effect of blocking oligonucleotide on predator/prey amplification

When amplifications were carried out only with *12SV5* primers, sequences of the leopard cat represented 91.6% of the total count, eight samples (sample 1–8; Fig. 3) exclusively yielded the leopard cat sequence, and 11 different prey taxa were observed in the diet. The blocking oligonucleotide *PrioB* drastically reduced the amplification of the leopard cat sequences, down to 2.2% of the total sequence count, with no leopard cat sequences observed in 31 samples. Under blocking nucleotide conditions, we recorded the amplification of seven additional prey items not previously detected when the same samples were amplified using the *12SV5* primers. The amplification failed in three samples when using the blocking oligonucleotide. The comparison of amplifications without and with blocking oligonucleotide is shown in Fig. 3.

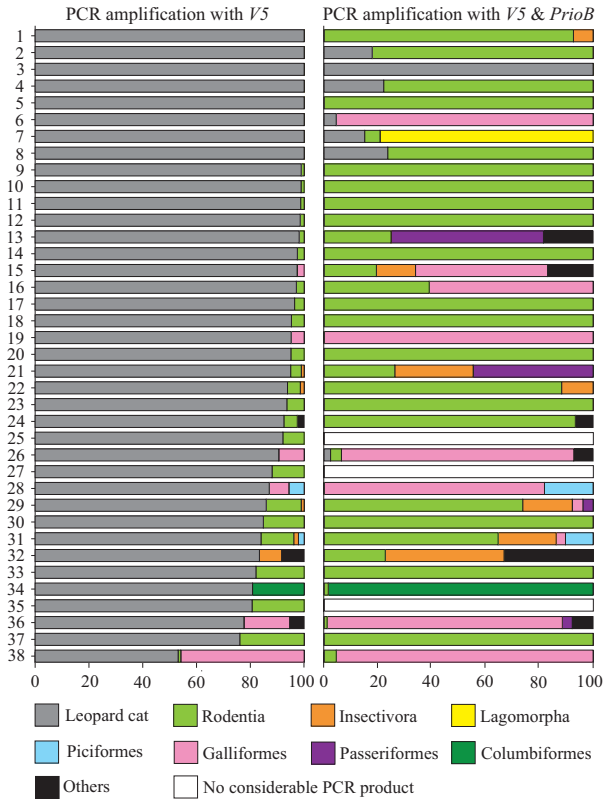


Fig. 3 Comparison of the amplifications of leopard cat and its prey sequences with *12SV5* primers for vertebrates without and with blocking oligonucleotide. The prey items are shown up to the order rank; fish and amphibians are grouped together in the ''others'' category. Each horizontal bar corresponds to the analysis of a single faeces using the *12SV5* primers, either without blocking oligonucleotide (on the left) or with blocking oligonucleotide (on the right). On each bar, the different colours represent the sequence count (%) of predator and prey items present in the sample. Samples 25, 27 and 35 did not show any considerable PCR products with blocking oligonucleotide amplification.

Table 3 Overview of the sequence counts at different stages of the analysis

Primer pair used	*12SV5F/12SV5R*	*12SV5F/12SV5R/PrioB*
Number of properly assembled sequences*	592 648	498 595
Number of unique sequences	44 441	73 414
Number of unique sequences, longer than 60 bp	44 066	46 765
Number of unique sequences, longer than 60 bp, with occurrence in the whole data set higher or equal to 100 (corresponding percentage of properly assembled sequences*)	232 (56.91%)	141 (44.84%)

*Direct and reverse sequence reads corresponding to a single DNA molecule were aligned and merged, producing what we called a ''properly assembled sequence''.

Diet composition of leopard cat

A total of 18 different prey taxa were identified in the diet of the leopard cat, seven of which were identified without ambiguity up to species level (Table 4). A maximum of seven prey items were observed within the same faeces sample, while 15 samples had only a single prey. We were not able to recover any prey DNA from only a single faeces: the experiments without and with blocking oligonucleotide with that sample produced only leopard cat sequences.

The diet composition of the leopard cat from ANP was eclectic; we observed 15 different prey taxa in 22 faeces samples. The house rat predominated the diet (in 68% of the faeces), followed by Asiatic white-toothed shrew (32%) and Murree hill frog (27%). We observed seven prey items (Himalayan wood mouse, Kashmir flying squirrel, Murree vole, Asiatic white-toothed shrew, chicken, kalij pheasant and jungle crow) within a single faeces, whereas six faeces indicated only a single prey. Overall, Rodentia dominates the diet at ANP with a presence in 91% of the faeces (Fig. 4a). Table 5 gives an overview of the leopard cat diet in Pakistan compared with previous studies.

Eight prey taxa were identified in 16 faeces from CGNP. The house rat predominated the diet (in 44% of the faeces), followed by Kashmir flying squirrel (31%) and Himalayan wood mouse (19%). Rodentia with five different prey species also dominated the diet at CGNP with a presence in 81% of the faeces (Fig. 4b).

While the leopard cat diet in both ANP and CGNP is composed mainly of rodents, the differences between these two areas were significant, both when considering all prey species independently (*P*-value: 0.01; χ^2 test with simulated *P*-value based on 10^6 replicates) and when grouping prey according to their taxonomy (Rodentia, Insectivora, Lagomorpha, Aves, Batracia and Teleostei; *P*-value: 0.03; χ^2 test with simulated *P*-value based on 10^6 replicates). As discussed in the study by Wright (2010), using Pearson chi-squared test for such data can lead to misinterpretation because of the overestimation of the degrees of freedom. By overestimating the degrees of freedom, it is more difficult to reject the null hypothesis. Consequently, rejecting the null hypothesis, as we did, places us on the conservative side of the decision.

Results of the rarefaction analysis are presented in Fig. 5. The number of prey species expected in the pooled faecal samples, based on the rarefaction curve, was 15 (95% CI: 13.91–16.09) and 8 (95% CI: 4.14–11.86) for the ANP and CGNP, respectively. In the case of ANP, 13 of 15 species with a cumulative frequency of 93% in the diet were detected in the first 11 samples. In CGNP, all of the documented prey species were detected in first 13 samples and the rest of the samples reflected their repeats.

Discussion

The leopard cat diet

All documented studies, including the present study, suggest that the order Rodentia is the primary prey base for the leopard cat (presence in 81.2–96.0% of the faeces in six studies, Table 5). Within *Rodentia*, the Muridae family dominates, with a presence in 50.0–86.4% of the faeces in Pakistan and up to 96% in other localities. The arboreal behaviour of the leopard cat (Nowell & Jackson 1996) broadens its trophic niche by enabling it to hunt tree-nesting birds and even flying squirrels in Pakistan. Birds and herpetofauna (reptiles and amphibians) are apparently the other main food groups after mammals. Birds have been reported in all studies, although the highest frequency was observed in Pakistan (presence in 18.7–45.5% of the faeces). In contrast to previous studies, where conventional methods did not allow species identification for birds, we are reporting eight distinct taxa. This specificity is an evident advantage of DNA-based diet methods recently developed. We also report fish in the diet, which have only once been reported previously (Inoue 1972). Our method did not allow the detection of invertebrates or plants, although these have been reported in other studies.

The results of the rarefaction analysis show the efficiency of the molecular method for detecting prey; this is advantageous for studying rare species that inhabit difficult terrains and that do not allow for collecting a large number of samples. Our sample size is smaller than what is generally recommended for classical diet studies; previously, 80 samples have been suggested for common leopards (Mukherjee *et al.* 1994). However, considering the greater detection efficiency of the new method, supported by the rarefaction estimates, our sample size seems to be adequate for estimating the vertebrate diet diversity of the leopard cat in the two studied regions.

The higher diversity of prey detected in samples from ANP as compared to those from CGNP probably reflects the higher productivity and diversity of temperate forests in the former park. The Kashmir flying squirrel prefers to nest on dead trees and is found in both national parks. Its frequency as a prey item was significantly higher in CGNP, the open forests of which probably make flying squirrel more susceptible to predation.

Surprisingly, the leopard cat seems to predate on prey with larger adult body size in Pakistan than in southern parts of its range (Table 5). Larger prey was

Table 4 List of prey taxa found in leopard cat diet (ANP: Ayubia National Park; CGNP: Chitral Gol National Park)

MOTU number	Accession number	Number of sequence reads	Number of occurrence		Most similar sequence(s) in public databases using BLAST				Putative taxon identification taking into account the locations where the faeces samples were collected	
			ANP 22 faeces	CGNP 16 faeces	Species name(s)	Accession number(s)	Query coverage (%)	Maximum identity (%)	Scientific name	Common name
1	FR873673	66680	15	7	*R. tanezumi/rattus*	EU273712/EU273707	100	100	*R. rattus*	House rat
2	FR873674	23746	4	0	*Microtus lusitanicus/ pyrenaicus/ duodecimcostatus/savii*	AJ972919/AJ972916/ AJ972915/AJ972914	100	95	*Hyperacrius wynnei* (?)	Murree vole (?)
3	FR873675	10848	4	0	*Phasianus colchicus/ versicolor*	FJ752430/AB164626	100	99	*Lophura leucomelanos* (?)	Kalij pheasant (?)
4	FR873676	10077	2	5	*Eoglaucomys fimbriatus*	AY227562	100	100	*E. fimbriatus*	Kashmir flying squirrel
5	FR873677	9902	5	3	*Apodemus uralensis*	AJ311128	100	100	*Apodemus rusiges*	Himalayan wood mouse
6	FR873678	9827	2	0	*Pucrasia macrolopha*	FJ752429	100	100	*P. macrolopha*	Koklass pheasant
7	FR873679	9361	7	0	*Crocidura gueldenstaedti*	AF434825	97	100	*Crocidura pullata* (?)	Asiatic white-toothed shrew (?)
8	FR873680	8700	0	1	*Columba livia*	GQ240309	100	99	*C. livia* (?)	Rock pigeon (?)
9	FR873681	8469	1	2	*Alectoris chukar*	FJ752426	100	100	*A. chukar*	Chukar partridge
10	FR873682	3626	6	0	*Nanorana parkeri*	AY322333	100	97	*Paa vicina* (?)	Murree hill frog (?)
11	FR873683	3329	0	1	*Lepus* spp.	AY292707	100	94	*Lepus capensis* (?)	Cape hare (?)
12	FR873684	2762	2	0	*Gallus gallus*	GU261719	100	100	*G. gallus*	Chicken
13	FR873687	2049	2	0	Timaliidae	AF376932	100	100	Timaliidae	Babblers
14	FR873688	1770	2	0	*Pica pica; Corvus macrorhynchos/corone/ frugilegus/albus*	HQ915867; AB042345/ AF386463/Y18522/ U38352	100	100	*C. macrorhynchos*	Jungle crow
15	FR873689	1034	2	0	*Picus viridis*	EF027325	100	97	*Dendrocopos* sp. (?)	Woodpecker (?)
16	FR873690	542	0	1	*Dryomys nitedula*	D89005	100	94	*D. nitedula* (?)	Forest dormouse (?)
17	FR873691	434	1	0	*Cephalosilurus apurensis; Liobagrus obesus*	EU179838; DQ321752	100	93	Siluriformes (?)	Cat fish (?)
18	FR873692	105	3	1	*Mus musculus castaneus*	EF108342	100	100	*Mus musculus*	House mouse

The question mark (?) after a taxon name means that the proposed identification must be confirmed by comparison with a local DNA database.

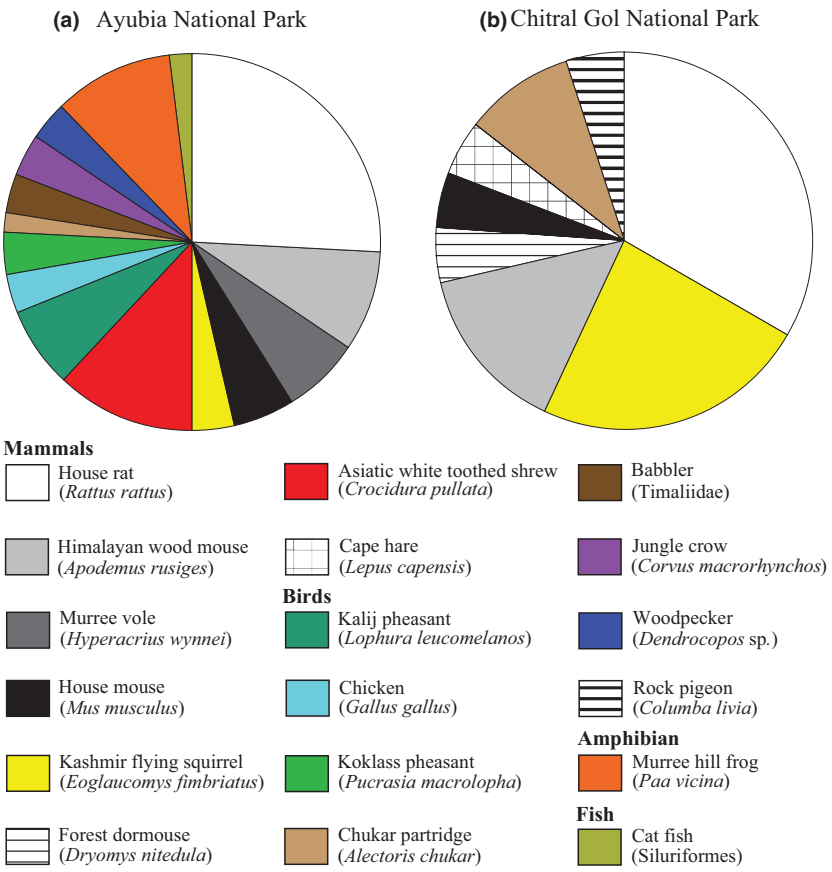


Fig. 4 Composition and comparison of the various prey items consumed and their relative frequency in the diet of the leopard cat at (a) Ayubia National Park and (b) Chitral Gol National Park.

Table 5 Comparison of leopard cat diet across its range in Asia. Except the present study, all other references estimated the diet using traditional morphology-based methods

	Occurrence in faeces, %							
Locality	Rodentia	Other mammals	Birds	Reptiles and amphibians	Fish	Invertebrates	Plant matter	References
ANP, Pakistan	90.9	31.8	45.5	27.3	4.5	Not recorded	Not recorded	Present study
CGNP, Pakistan	81.2	6.2	18.7	0.0	0.0	Not recorded	Not recorded	Present study
Negros-Panay Faunal Region, Philippines	96.0	8.0	8.0	—	—	—	12.0	Fernandez & de Guia (2011)
Khao Yai National Park, Thailand	93.8	24.5	8.2	8.2	—	36.7	—	Austin *et al.* (2007)
Sabah, Malaysian Borneo	93.1	4.2	5.6	19.4	—	11.1	11.1	Rajaratnam *et al.* (2007)
North-central Thailand	89.0	17.0	4.0	—	—	21.0	—	Grassman *et al.* (2005)
Tsushima islands, Japan	91.3	0.3	36.5	22.3	—	24.3	78.8	Tatara & Doi (1994)

usually the house rat (140–280 g), but even bigger prey were occasionally reported. Grassman *et al.* (2005) found remains of Java mouse deer (*Tragulus javanicus*; 1.18–1.28 kg from Weathers & Snyder (1977) and Endo *et al.* 2002) in leopard cat faeces, and Austin *et al.* (2007) once recorded a large ungulate (*Cervus unicolor*; 70.5–112 kg from Idris *et al.* 2000). In Pakistan, many large prey were found in the diet, including the Kashmir flying squirrel (560–734 g; Hayssen 2008), the cape hare (2.10–2.30 kg; Lu 2000), the chukar partridge (450–800 g; del Hoyo *et al.* 1994), the kalij pheasant (564–1150 g; del Hoyo *et al.* 1994), the koklass pheasant

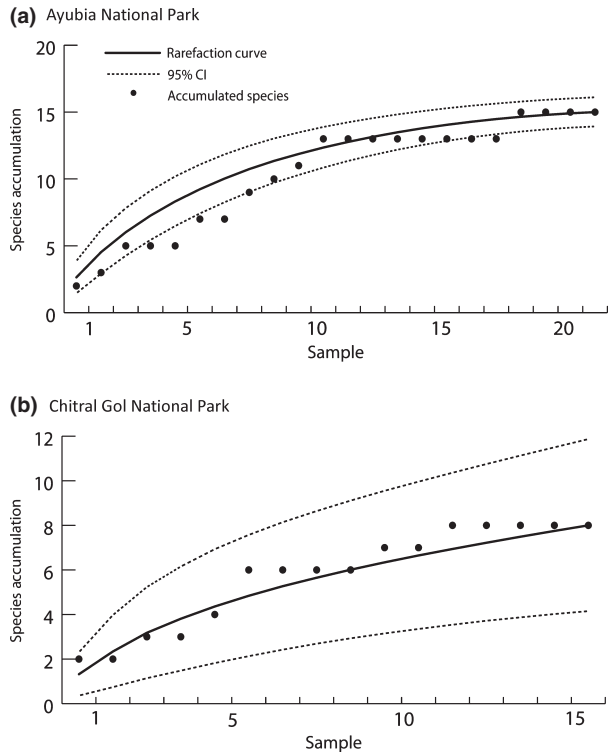


Fig. 5 Species accumulation curves based on the prey species identified in the faecal samples of leopard cat collected in ANP (a) and CGNP (b).

(930–1415 g; del Hoyo *et al.* 1994) and the jungle crow (570–580 g; Matsubara 2003). Two nonexclusive explanations can be proposed to explain such a diet shift towards larger species. First, only juveniles of the larger species may have been captured. It is important to note that remains of juveniles might be difficult to identify in faeces using traditional approaches. DNA-based methods allow straightforward taxon identification, but obviously not the age of prey. Second, the body size of the leopard cat in Pakistan might be larger than in southern areas of its distribution range, possibly explaining their ability to catch larger prey. This last hypothesis tends to be supported by the fact that the leopard cat is known to show considerable variation in size across its geographic distribution, with larger animals in China and Russia (Sunquist & Sunquist 2009), but cannot be confirmed because of the scarcity of data in Pakistan.

We conclude that the results of the present study are in general agreement with previous diet studies of the leopard cat indicating a very eclectic diet. However, the present study highlighted a possible broadening of the diet to include larger prey and provided more precise information by resolving major diet groups to a lower taxonomical level, which was not previously possible using conventional methods.

Conservation implications

The current extent of occurrence of the leopard cat in Pakistan is not resolved (Sheikh & Molur 2004). It historic range started from Chitral and extended to the eastern border of Pakistan, including areas of Swat, Hazara and Ayubia National Park (Nowell & Jackson 1996; Roberts 2005). In the north, it occupied parts of Gilgit Baltistan probably up to an elevation of 3000 m (Habibi 1977). The present study documents its current occurrence in two extremities of its historic range. A leopard cat was photographed in CGNP (SLT 2008), and authors have collected evidence of its presence in Machiara National Park, Azad Jammu and Kashmir, and western parts of the Gilgit Baltistan. This evidence suggests that the historic range of the cat in Pakistan is probably intact, although its population status needs to be determined.

Among the 18 taxa eaten by the cat in Pakistan, four (*Apodemus rusiges*, *Dryomys nitedula*, *Eoglaucomys fimbriatus* and *Lepus capensis*) are categorized as vulnerable (Sheikh & Molur 2004). Because the leopard cat is highly adaptable and appears to be widespread in Pakistan, it may be a potential threat to these species, which have a cumulative frequency of 44.7% of occurrence in faeces. A population assessment of the leopard cats is needed to evaluate the magnitude of this possible threat and to tailor an appropriate management strategy for both prey and predator.

A DNA-based approach for studying carnivore diet

Diet analysis combining next-generation sequencing and vertebrate primers with blocking oligonucleotides has tremendous potential for large-scale studies on carnivore diet. This approach is very robust and presents the complete diet profile of the vertebrate prey consumed. It is highly accurate and discriminates between closely related species in most of the cases. Moreover, a priori knowledge of prey items consumed is not essential, as it is when designing more specific DNA-based approaches. However, such analyses can yield a substantial amount of artefactual sequences including chimeras, nuclear pseudogenes and primer dimers, especially when using the blocking oligonucleotide. As our primers target highly conserved DNA regions in vertebrates, it seems unlikely that a nuclear pseudogene will better match with the *12SV5* primers than the true mitochondrial copies. Furthermore, as mitochondrial copies are much more frequent than nuclear copies, the number of occurrences of any pseudogene sequence should be much lower than the corresponding mitochondrial sequence. With regard to these possible artefacts, we recommend keeping stringent PCR conditions

as described in the Materials and Methods section and treating as significant only sequences showing a strong correspondence with a known sequence (at least >0.9) together with a relatively high frequency.

An ongoing debate on DNA-based diet studies concerns the quantification of different prey items consumed and their relative presence in sequence counts. This issue has been highlighted in several recent DNA-based dietary studies (e.g. Deagle *et al.* 2009, 2010; Soininen *et al.* 2009; Valentini *et al.* 2009a). The sequence count cannot be interpreted as quantitative for a few reasons. Biased amplification of some species has been observed when PCR was carried out of a known mixture (Polz & Cavanaugh 1998). Strong biases will occur in dietary studies when primers mismatch with certain prey sequences, resulting in the amplification inclined towards the perfect matches. The two highly conserved regions targeted by the primers *12SV5F* and *12SV5R* make the approach less susceptible to PCR biases. Deagle *et al.* (2010) suggested that differences in the density of mitochondrial DNA in tissues can also bias the sequence count. In the present study, we avoided quantitative interpretations from the results of our sequence counts and recorded only the presence/absence of the different prey in the different faeces.

The blocking oligonucleotide approach has considerable potential for its use in trophic analyses. The design of a blocking oligonucleotide specific to the leopard cat requires knowing the leopard cat sequence for the target DNA region. In this study, the blocking oligonucleotide technique not only inhibited the amplification of the leopard cat DNA, but also uncovered seven more prey taxa in the diet that had not been amplified previously without the blocking oligonucleotide. We used a high concentration of *PrioB* (2 μM) compared with *12SV5F* and *12SV5R* primers (0.1 μM each). For each faeces sample, we systematically ran amplifications without and with blocking oligonucleotide, as amplification with such a relatively high *PrioB* concentration might fail.

One limitation of the approach with the *12SV5F* and *12SV5R* primers proposed here is that it only identifies vertebrate prey. Many carnivores have a more diverse diet, including invertebrates and plants. For example, the Eurasian badger (*Meles meles*) exploits a wide range of food items, especially earthworms, insects and grubs. It also eats small mammals, amphibians, reptiles and birds as well as roots and fruits (Revilla & Palomares 2002). For instance, to study the badger's diet, we suggest complementing the primers for vertebrates with several additional systems, such as primers targeting plant taxa (e.g. Taberlet *et al.* 2007; Valentini *et al.* 2009a) or earthworms (Bienert *et al.* 2012).

One more limitation of this approach for identifying vertebrates is that cases of cannibalism cannot be detected. In such a situation, the predator DNA cannot be distinguished from the prey DNA that belongs to the same species. This limitation was not acknowledged in previous DNA-based diet analyses for vertebrate predators, despite the cases of cannibalism have been documented, for example, in Otariidae (e.g. Wilkinson *et al.* 2000). However, if cannibalism is important from a behavioural point of view, it represents a marginal phenomenon when studying the diet.

Another potential difficulty concerns species identification. In some cases, we had to combine the best match using public databases together with expert knowledge about the available prey in the location where the faeces were collected. For example, in our study, the best match (99%) for MOTU number 3 in public databases corresponded to two species of the genus *Phasianus* (*P. colchicus* and *P. versicolor*). These two species are not recorded in ANP, and thus, we identify this MOTU as the closest relative (Huang *et al.* 2009; Shen *et al.* 2010) occurring in ANP, the kalij pheasant (*Lophura leucomelanos*). If the identification of the kalij pheasant seems reliable, some other putative identifications are more problematic, particularly those having a relatively low identities with known sequences in public databases (i.e. *Hyperacrius wynnei, Paa vicina, L. capensis, Dendrocopos* sp., and *D. nitedula*). To remove such uncertainties, we recommend constructing a local reference database when possible.

The results of the present study correspond to summer diet and may not reflect the complete diet profile of the leopard cat in Pakistan. In future, it would be interesting to collect samples throughout the year, with the attendant possibility of revealing more prey taxa than what we have observed in this study.

Conclusion

Noninvasive sampling is the only way to study the diet of elusive animals like the leopard cat. In Pakistan, we obtained results confirming the eclectic characteristics of this predator, together with an extension of the diet towards larger prey. The DNA-based approach has a better resolution than conventional approach-based identification of prey from hair and bone remains. While DNA-based methods cannot assess prey ages, conventional approaches might reveal the potential ages of the prey when necessary, possibly determining whether juveniles or adults of larger prey were consumed. As a consequence, DNA-based diet analysis can provide a valuable complement to conventional methods.

The DNA-based approach we propose here is particularly robust and simple to implement and allows the possibility of very large-scale analyses. It can be applied

to other carnivore species with only a slight adjustment concerning the design of the blocking oligonucleotide.

Acknowledgements

The authors are thankful to A. Valentini, D. Rioux and L. Gielly for their valuable suggestions and comments throughout the course of this study. For English correction in the manuscript, we express our gratitude to M. MacFarlane and M. Reynolds. W. Shehzad and T. Riaz were supported by PhD scholarship from the HEC (Higher Education Commission), Government of Pakistan. The Himalayan Wildlife Foundation and Snow Leopard Trust provided financial support for sample collections in Ayubia and Chitral Gol National Parks. The staff of the Khyber Pakhtunkhwa Wildlife Department and Snow Leopard Trust assisted in sample collection.

References

Austin SC, Tewes ME, Grassman Jr LI, Silvy NJ (2007) Ecology and conservation of the leopard can *Prionailurus bengalensis* and clouded leopard *Neofelis nebulosa* in Khao Yai National Park, Thailand. *Acta Zoologica Sinica*, **53**, 1–14.

Bagchi S, Mishra C (2006) Living with large carnivores: predation on livestock by the snow leopard (*Uncia uncia*). *Journal of Zoology*, **268**, 217–224.

Bellemain E, Carlsen T, Brochmann C *et al*. (2010) ITS as DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiology*, **10**, 189.

Bienert R, de Danieli S, Miquel C *et al*. (2012) Tracking earthworm communities from soil DNA. *Molecular Ecology*, **21**, doi: 10.1111/j.1365-294X.2011.05407.x.

Blankenship LE, Yayanos AA (2005) Universal primers and PCR of gut contents to study marine invertebrate diets. *Molecular Ecology*, **14**, 891–899.

Bradley BJ, Stiller M, Doran-Sheehy DM *et al*. (2007) Plant DNA sequences from feces: potential means for assessing diets of wild primates. *American Journal of Primatology*, **69**, 699–705.

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* polymerase: primer modification that facilitate genotyping. *BioTechniques*, **20**, 1004–1010.

Colwell RK, Mao CX, Chang J (2004) Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology*, **85**, 2717–2727.

Deagle BE, Jarman SN, Pemberton D, Gales NJ (2005a) Genetic screening for prey in the gut contents from a giant squid (*Architeuthis* sp.). *Journal of Heredity*, **96**, 417–423.

Deagle BE, Tollit DJ, Jarman SN *et al*. (2005b) Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. *Molecular Ecology*, **14**, 1831–1842.

Deagle BE, Eveson JP, Jarman SN (2006) Quantification of damage in DNA recovered from highly degraded samples – a case study on DNA in faeces. *Frontier in Zoology*, **3**, 11.

Deagle BE, Gales NJ, Evans K *et al*. (2007) Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (*Eudyptes chrysolophus*). *PLoS ONE*, **2**, e831.

Deagle BE, Kirkwood R, Jarman SN (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, **18**, 2022–2038.

Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics*, **11**, 2039–2048.

Dunshea G (2009) DNA-based diet analysis for any predator. *PLoS ONE*, **4**, e5252.

Endo H, Kimura J, Sasaki M *et al*. (2002) Functional morphology of the mastication muscles in the lesser and greater mouse deer. *Journal of Veterinary Medical Science*, **64**, 901–905.

Farooque M (2007) *Management Plan of Ayubia National Park*. Khyber Pakhtunkhwa Wildlife Department, Peshawar.

Fernandez DAP, de Guia APO (2011) Feeding habits of Visayan leopard cats (*Prionailurus bengalensis rabori*) in sugarcane fields of Negros Occidental, Philippines. *Asia Life Sciences*, **20**, 143–154.

Ficetola GF, Coissac E, Zundel S *et al*. (2010) An *in silico* approach for the evaluation of DNA barcodes. *BMC Genomics*, **11**, 434.

Gaines WL (2001) Large carnivore surveys in the North Karakorum Mountains, Pakistan. *Natural Areas Journal*, **21**, 168–171.

GoNWFP & IUCN (1996) *Sarhad Provincial Conservation Strategy*. Sarhad Programme Office, IUCN–The World Conservation Union, Peshawar.

Grassman LI, Tewes ME, Silvy NJ, Kreetiyutanont K (2005) Spatial organization and diet of the leopard cat (*Prionailurus bengalensis*) in north-central Thailand. *Journal of Zoology*, **266**, 45–54.

Green SJ, Minz D (2005) Suicide polymerase endonuclease restriction, a novel technique for enhancing PCR amplification of minor DNA templates. *Applied and Environmental Microbiology*, **71**, 4721–4727.

Habibi K (1977) *The Mammals of Afghanistan: Their Distribution and Status*. UNDP, FAO and Ministry of Agriculture, Kabul.

Hayssen V (2008) Patterns of body and tail length and body mass in Sciuridae. *Journal of Mammalogy*, **89**, 852–873.

Höss M, Kohn M, Pääbo S, Knauer F, Schröder W (1992) Excrement analysis by PCR. *Nature*, **359**, 199.

del Hoyo J, Elliot A, Sargatal J (1994) *Handbook of the Birds of the World*, Vol. 2. New World Vultures to Guinea Fowl Lynx Edicions, Barcelona.

Huang ZH, Liu NF, Xiao YA *et al*. (2009) Phylogenetic relationships of four endemic genera of the Phasianidae in China based on mitochondrial DNA control-region genes. *Molecular Phylogenetics and Evolution*, **53**, 378–383.

Idris I, Moin S, Sulah S, Jiwan D (2000) Some physical characteristics of Sambar deer (*Cervus unicolor*). *Pertanika Journal of Tropical Agricultural Science*, **33**, 55–59.

Inoue T (1972) The food habit of Tsushima leopard cat, *Felis bengalensis* spp., analysed from their scats. *Journal of the Mammalogical Society of Japan*, **5**, 155–169.

Izawa M, Doi T (1991) Status of conservation and management of two species of felidae in Japan (in Japanese). *Mammalian Science*, **31**, 15–22.

Jarman SN, Gales NJ, Tierney M, Gill PC, Elliott NG (2002) A DNA-based method for identification of krill species and its

application to analysing the diet of marine vertebrate predators. *Molecular Ecology*, **11**, 2679–2690.

Jarman SN, Deagle BE, Gales NJ (2004) Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology*, **13**, 1313–1322.

Jarman SN, Redd KS, Gales NJ (2006) Group-specific primers for amplifying DNA sequences that identify Amphipoda, Cephalopoda, Echinodermata, Gastropoda, Isopoda, Ostracoda and Thoracica. *Molecular Ecology Notes*, **6**, 268–271.

Kageyama T, Sato Y, Nishizawa S, Teramae N (2008) Competitive binding of small ligands to nucleobases in AP site-containing DNA duplexes. *Nucleic Acids Symposium Series*, **52**, 119–120.

King RA, Vaughan IP, Bell JR, Bohan DA, Symondson WOC (2010) Prey choice by carabid beetles feeding on an earthworm community analysed using species- and lineage-specific PCR primers. *Molecular Ecology*, **19**, 1721–1732.

Kohn MH, Wayne RK (1997) Facts from feces revisited. *Trends in Ecology & Evolution*, **12**, 223–227.

Krahn MM, Herman DP, Matkin CO *et al.* (2007) Use of chemical tracers in assessing the diet and foraging regions of eastern North Pacific killer whales. *Marine Environmental Research*, **63**, 91–114.

Lau MWN, Fellowes JR, Chan BPL (2010) Carnivores (Mammalia: Carnivora) in South China: a status review with notes on the commercial trade. *Mammal Review*, **40**, 247–292.

Li MJ, Sato Y, Nishizawa S *et al.* (2009) 2-aminopurine-modified abasic-site-containing duplex DNA for highly selective detection of theophylline. *Journal of the American Chemical Society*, **131**, 2448–2449.

Liles MR, Manske BF, Bintrim SB, Handelsman J, Goodman RM (2003) A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Applied and Environmental Microbiology*, **69**, 2684–2691.

Lovari S, Boesi R, Minder I *et al.* (2009) Restoring a keystone predator may endanger a prey species in a human-altered ecosystem: the return of the snow leopard to Sagarmatha National Park. *Animal Conservation*, **12**, 559–570.

Lu X (2000) Body weights of the Cape hare *Lepus capensis* in the northern China. *Acta Theriologica*, **45**, 271–280.

Macdonald DW, Loveridge AJ, Nowell K (2010) Dramatis personae: an introduction to the wild felids. In: *Biology and Conservation of Wild Felids* (eds Macdonald DW and Loveridge AJ), pp. 3–58. Oxford University Press, New York.

Magnuson VL, Ally DS, Nylund SJ *et al.* (1996) Substrate nucleotide-determinated non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. *BioTechniques*, **21**, 700–709.

Matsubara H (2003) Comparative study of territoriality and habitat use in syntopic Jungle Crow (*Corvus macrorhynchos*) and Carrion Crow (*C. corone*). *Ornithological Science*, **2**, 103–111.

Mills LS, Soule ME, Doak DF (1993) The keystone-species concept in ecology and conservation. *BioScience*, **43**, 219–224.

Mirza ZB (2003) *Biological Baseline Study of Chitral Gol National Park*. Protected Areas Management Project, Islamabad.

Mitani N, Mihara S, Ishii N, Koike H (2009) Clues to the cause of the Tsushima leopard cat (*Prionailurus bengalensis euptilura*) decline from isotopic measurements in three species of Carnivora. *Ecological Research*, **24**, 897–908.

Mukherjee S, Goyal SP, Chellam R (1994) Standardization of scat analysis techniques for leopard (*Panthera pardus*) in Gir National Park, western India. *Mammalia*, **58**, 139–143.

Mukherjee S, Krishnan A, Tamma K *et al.* (2010) Ecology driving genetic variation: a comparative phylogeography of jungle cat (*Felis chaus*) and leopard cat (*Prionailurus bengalensis*) in India. *PLoS ONE*, **5**, 16.

Natoli E (1990) Mating strategies in cats – a comparison of the role and importance of infanticide in domestic cats, *Felis catus* L and lions, *Panthera leo* L. *Animal Behaviour*, **40**, 183–186.

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, **48**, 443–453.

Nowell K, Jackson P (1996) *Wild Cats, Status Survey and Conservation Action Plan*. IUCN, Gland, Switzerland, 382p.

Oli MK, Taylor IR, Rogers ME (1994) Snow leopard (*Panthera Uncia*) predation of livestock – an assessment of local perceptions in the Annapurna Conservation area, Nepal. *Biological Conservation*, **68**, 63–68.

Pegard A, Miquel C, Valentini A *et al.* (2009) Universal DNA based methods for assessing the diet of grazing livestock and wildlife from faeces. *Journal of Agricultural and Food Chemistry*, **57**, 5700–5706.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, **64**, 3724–3730.

Power ME, Tilman D, Estes JA *et al.* (1996) Challenges in the quest for keystones. *BioScience*, **46**, 609–620.

Rajaratnam R, Sunquist M, Rajaratnam L, Ambu L (2007) Diet and habitat selection of the leopard cat (*Prionailurus bengalensis borneoensis*) in an agricultural landscape in Sabah, Malaysian Borneo. *Journal of Tropical Ecology*, **23**, 209–217.

Revilla E, Palomares F (2002) Does local feeding specialization exist in Eurasian badgers? *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **80**, 83–93.

Rho P (2009) Use of GIS to develop a multivariate habitat model for the leopard cat (*Prionailurus bengalensis*) in mountainous region of Korea. *Journal of Ecology & Field Biology*, **32**, 229–236.

Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, **39**, e145.

Roberts TJ (2005) *Field Guide to the Large and Medium Sized Mammals of Pakistan*. Oxford University Press, Oxford, UK.

Scott DM, Gemita E, Maddox TM (2004) Small cats in human modified landscapes in Sumatra. *Cat News*, **40**, 23–25.

Sheikh KM, Molur S (2004) Status and red list of Pakistan's mammals. Based on the Conservation Assessment and Management Plan Workshop. IUCN Pakistan.

Shen YY, Liang L, Sun YB *et al.* (2010) A mitogenomic perspective on the ancient, rapid radiation in the Galliformes with an emphasis on the Phasianidae. *BMC Evolutionary Biology*, **10**, 132.

Snow Leopard Trust (2008) Population monitoring of large carnivores in Chitral Gol National Park. Progress update for 2007–2008, submitted to NWFP Wildlife Department, Peshawar, Pakistan.

Soininen EM, Valentini A, Coissac E *et al.* (2009) Analysing diet of small herbivores: the efficiency of DNA barcoding

coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in Zoology*, **6**, 9.

Sunquist ME, Sunquist FC (2002) *Wild Cats of the World*. University of Chicago Press, Chicago, Illinois, 416pp.

Sunquist ME, Sunquist FC (2009) Family Felidae (cats). In: *Handbook of the Mammals of the World. Vol. 1. Carnivores* (eds Wilson DE and Mittermeier RA), pp. 54–168. Lynx Edicions, Barcelona.

Symondson WOC (2002) Molecular identification of prey in predator diets. *Molecular Ecology*, **11**, 627–641.

Taberlet P, Coissac E, Pompanon F *et al.* (2007) Power and limitations of the chloroplast *trn*L(UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, **35**, e14.

Tatara M, Doi T (1994) Comparative analyses on food-habits of Japanese marten, Siberian weasel and leopard cat in The Tsushima islands, Japan. *Ecological Research*, **9**, 99–107.

Valentini A, Miquel C, Nawaz MA *et al.* (2009a) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trn*L approach. *Molecular Ecology Resources*, **9**, 51–60.

Valentini A, Pompanon F, Taberlet P (2009b) DNA barcoding for ecologists. *Trends in Ecology and Evolution*, **24**, 110–117.

Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology*, **5**, 11.

Vestheim H, Edvardsen B, Kaartvedt S (2005) Assessing feeding of a carnivorous copepod using species-specific PCR. *Marine Biology*, **147**, 381–385.

Wang YF, Ng MTT, Zhou TY *et al.* (2008) C3-Spacer-containing circular oligonucleotides as inhibitors of human topoisomerase I. *Bioorganic & Medicinal Chemistry Letters*, **18**, 3597–3602.

Watanabe S (2009) Factors affecting the distribution of the leopard cat *Prionailurus bengalensis* on East Asian islands. *Mammal Study*, **34**, 201–207.

Weathers WW, Snyder GK (1977) Hemodynamics of the lesser mouse deer, *Tragulus javanicus*. *Journal of Applied Physiology*, **42**, 679–681.

Wilkinson IS, Childerhouse SJ, Duignan PJ, Gulland FMD (2000) Infanticide and cannibalism in the New Zealand sea lion, *Phocarctos hookeri*. *Marine Mammal Science*, **16**, 494–500.

Wright BE (2010) Use of chi-square tests to analyze scat-derived diet composition data. *Marine Mammal Science*, **26**, 395–401.

Zaidi RH, Jaal Z, Hawkes NJ, Hemingway J, Symondson WOC (1999) Can multiple-copy sequences of prey DNA be detected amongst the gut contents of invertebrate predators? *Molecular Ecology*, **8**, 2081–2087.

T.R., E.C. and P.T. are co-inventors of a pending French patent on the primer pair named *12SV5F* and *12SV5R* and on the use of the amplified fragment for identifying vertebrate species from environmental samples. This patent only restricts commercial applications and has no impact on the use of this method by academic researchers.

Data accessibility

DNA sequences of the V5 loop of the mitochondrial 12S gene: GenBank accessions FR873673–FR873692.

Fasta file and filtered data deposited in the Dryad repository: doi: 10.5061/dryad.443t4m1q.

4. Jungle without prey: livestock sustains population of an endangered cat in Pakistan

WASIM SHEHZAD, MUHAMMAD ALI NAWAZ, FRANÇOIS POMPANON, ERIC COISSAC, TIAYYBA RIAZ, SAFDAR ALI SHAH and PIERRE TABERLET

WASIM SHEHZAD, FRANÇOIS POMPANON, ERIC COISSAC, TIAYYBA RIAZ, PIERRE TABERLET (Corresponding author) Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 9, France.
E-mail pierre.taberlet@ujf-grenoble.fr,
MUHAMMAD ALI NAWAZ Snow Leopard Trust (Pakistan Program), 17-Service Road North, I-8/3, Islamabad, Pakistan
SAFDAR ALI SHAH Wildlife Department, Khyber Pakhtunkhwa, Pakistan

Received:

Corresponding author: Pierre Taberlet, Laboratoire d'Ecologie Alpine, CNRS-UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 09, France; Tel: +33(0)4 76 63 54 33; Fax: +33(0)4 76 51 42 79; E-mail: pierre.taberlet@ujf-grenoble.fr

Word count: 5933 words

Study area and sampling

Ayubia National Park was established in 1984 and is spread over 33 square kilometers; surrounding reserved forests cover an area of 150 square kilometers (Fig. 1). The park is comprised of moist temperate forests, sub-alpine meadows, and sub-tropical pine forests. The altitudinal variation ranges from 1050 - 3027 m a.s.l., and it receives a mean annual rainfall between 1065 - 1424mm. It has approximately 200 species of birds, 31 species of mammals, 16 species of reptiles and three species of amphibians (Farooque, 2007). Three historically occurring wild ungulate species (Musk deer *Moschus moschiferus*, Grey goral *Naemorhedus goral*, Barking deer *Muntiacus muntjak*) have been illegally hunted to extirpation from the area in the recent past. The leopard population had become extremely rare by the early 1980s and was close to extinction in Galiat region and adjacent areas. However, the creation of ANP in 1984 served as a healthy means to recover some species like the common leopard and Rhesus monkey (*Macaca mulatta*) (Lodhi, 2007).

One hundred and eleven putative common leopard fecal samples were collected in ANP during the summer season in 2008. These samples were preserved first in 90 % ethanol and then shifted into silica gel for transportation to LECA (Laboratoire d'Ecologie Alpine), Université Joseph Fourier, Grenoble, France, for diet analysis of the common leopard.

Methods

DNA extraction

All extractions were performed in a room dedicated to the extraction of degraded DNA. Total DNA was extracted from about 15 mg of feces with the DNeasy Blood and Tissue Kit (QIAgen GmbH, Hilden, Germany), following the manufacturer's instructions, with a slight modification at the beginning of the protocol as described by Shehzad et al. (submitted). The DNA extracts were recovered in a total volume of 250 µL. Blank extractions without samples were systematically performed to monitor possible contaminations.

Primer design

Identification of predator species

A primer pair *PantF/PantR* highly specific to the common leopard was designed on the 12S mitochondrial rRNA gene (Table 1), with the 3'-end of each primer as different as possible from other species. The specificity of this primer pair, amplifying a 79 bp fragment, was validated in silico by using the *ecoPCR* program (Bellemain et al., 2010; Ficetola et al., 2010), with the following parameters: a perfect match on the two last nucleotides, and a maximum of three mismatches on the remaining nucleotides. With these parameters, only the *P. pardus* mitochondrial 12S gene was recovered. This in silico validation confirmed that *PantF/PantR* should unambiguously identify common leopard feces. The PCRs were carried out in a total volume of 20 µl with 8 mM Tris-HCl (PH 8.3), 40 mM KCl, 2 mM $MgCl_2$, 0.2 µM of each primer, BSA (5 µg), 0.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) using 2 µl as DNA template. The PCR conditions were chosen as an initial 10 min denaturation step at 95°C, followed by 45 cycles of denaturation at 95°C for 30 s and annealing at 53°C for 30 s. Thus, the primary identification of the samples was done on the basis of the presence of a PCR product of the suitable length revealed by electrophoresis on a 2 % agarose gel. The samples successfully amplified by the *PantF/PantR* primer pair were selected for further analyses.

Universal primer pair for vertebrates *12SV5F/12SV5R*

We used the primer pair *12SV5F*/*12SV5R* designed by the *ecoPrimers* program (Riaz et al., 2011). The *ecoPrimers* scans whole genomes to find new barcode markers and their associated primers. This program optimizes two quality indices measuring the taxonomical coverage and the discrimination power to select the most efficient markers,

according to specific experimental constraints such as marker length or targeted taxa. This universal primer pair for vertebrates represents the best choice found by *ecoPrimers* among short barcodes according to the available vertebrate whole mitochondrial genomes currently available. It amplifies a ~100 bp fragment of the V5 loop of the mitochondrial 12S ribosomal gene, with the ability to amplify short DNA fragments such as recovered from feces, and has a high taxonomic resolution despite its short size. Using the ecoPCR program (Bellemain et al., 2010; Ficetola et al., 2010), and based on the release 103 of the EMBL database, this fragment unambiguously identifies 77 % of the species and 89 % of the genera.

Blocking oligonucleotide specific to leopard sequence

The *PantB* (Table 1) blocking oligonucleotide specific to leopard sequence was designed as suggested by Vestheim & Jarman (2008), and overlaps the amplification primer *12SV5R* by six nucleotides. This blocking oligonucleotide was used to restrain the amplification of common leopard sequences when using the universal primers targeting all vertebrate prey.

DNA amplification for diet analysis

All DNA amplifications were carried out in a final volume of 25 µL, using 2 µL of DNA extract as a template. The amplification mixture contained 1 U of AmpliTaq® Gold DNA Polymerase (Applied Biosystems, Foster City, CA), 10 mM Tris-HCl, 50 mM KCl, 2 mM of MgCl2, 0.2 mM of each dNTP, 0.1 µM of each primer (*12SV5F/12SV5R*) with 2 µM of *PantB*, and 5 µg of bovine serum albumin (BSA, Roche Diagnostic, Basel, Switzerland). The PCR mixture was denatured at 95°C for 10 min, followed by 45 cycles of 30 s at 95°C, and 30 s at 60°C; as the target sequences are about 100 bp long, the elongation step was removed to reduce the +A artifact (Brownstein et al., 1996; Magnuson et al., 1996) that might decrease the efficiency of the first step of the sequencing process (blunt-end ligation).

The universal primers for vertebrates *12SV5* were modified by the addition of specific tags on the 5′ end to allow the assignment of sequence reads to the relevant sample (Valentini et al., 2009). As a consequence, all the PCR products were tagged identically on both ends. These tags were composed of CC on the 5' end followed by seven variable nucleotides that were specific to each sample. The seven variable nucleotides were designed using the *oligoTag* program (www.prabi.grenoble.fr/trac/OBITools) with at least three differences among the tags,

without homopolymers longer than two, and avoiding a C on the 5' end. All the PCR products from the different samples were first purified using the MinElute PCR purification kit (QIAGEN GmbH) then titrated using capillary electrophoresis (QIAxel, QIAgen GmbH, Hilden, Germany) and finally mixed together, in equimolar concentration, before the sequencing.

DNA sequencing

The sequencing was carried out on the Illumina/Solexa Genome Analyzer IIx (Illumina Inc., San Diego, CA 92121 USA), using the Paired-End Cluster Generation Kit V4 and the Sequencing Kit V4 (Illumina Inc., San Diego, CA 92121 USA), and following manufacturer's instructions. A total of 108 nucleotides were sequenced on each extremity of the DNA fragments.

Sequence analysis and taxon assignation

The sequence reads were analyzed using the OBITools (www.prabi.grenoble.fr/trac/OBITools). First, the direct and reverse reads corresponding to a single molecule were aligned and merged using the *solexaPairEnd* program, taking into account data quality during the alignment and the consensus computation. Then, primers and tags were identified using the *ngsfilter* program. Only sequences with a perfect match on tags and a maximum of two errors on primers were taken into account. The amplified regions, excluding primers and tags, were kept for further analysis. Strictly identical sequences were clustered together using the *obiuniq* program, keeping the information about their distribution among samples. Sequences shorter than 60 bp, or containing ambiguous nucleotides, or with occurrences lower or equal to 100 were excluded using the *obigrep* program. The *obiclean* program was then implemented to detect amplification/sequencing errors, by giving each sequence within a PCR product the status of "head" (most common sequence among all sequences that can be linked with a single indel or subsitution), "singleton" (no other variant with a single difference in the relevant PCR product), or "internal" (all other sequences not being "head" or "singleton", i.e. corresponding to amplification/sequencing errors). Taxon assignation was achieved using the *ecoTag* program (Pegard et al., 2009). *EcoTag* relies on the FASTA35 program (Needleman & Wunsch, 1970) to find highly similar sequences in the reference database. This database was built by extracting the relevant part of the mitochondrial 12S gene from EMBL nucleotide library (release 107) using the ecoPCR program (Ficetola et al., 2010). A unique taxon was assigned to each unique sequence.

This unique taxon corresponds to the last common ancestor node in the NCBI taxonomic tree of all the taxids, annotating the sequences of the reference database that matched against the query sequence. A final filtering was carried out by removing sequences that were never "head" or "singleton", sequences that were not identified at the family level (for removing putative chimeras), and the sequences with a total count among the whole dataset inferior to 750 (plus the removal of an obvious human contamination). Finally, automatically assigned taxonomic identification was completed manually by combining the automatic identification with distribution data of prey in the study area to obtain a more precise identification when relevant.

Results

Of 111 presumed common leopard fecal samples received from ANP, 60 (54.0 %) samples were identified as belonging to this species using *PantF/PantR* specific primers and selected for further experimentation. After assembling the forward and reverse reads, and filtering for primers and tags, we obtained a total of 652 090 sequences, corresponding to 36 929 unique sequences (DRYAD entry: doi:XXXXXX). The removal of sequences shorter than 60 bp and with a total count of less than 100 reduced the dataset to 150 sequences. Ninety-four sequences that were never "head" or "singleton" were removed, as were 41 sequences with a total count among all samples of less than 750. This threshold has been determined in order to avoid, having twice the same species identification in the result list, and with a lower identity for sequences below the threshold. Finally, three sequences without a perfect identity corresponding to the reference database were also removed, leading to 12 (including one *Panther pardus* plus 11 prey sequences) MOTU (Molecular Operational Taxonomic Unit; Blaxter, 2004). The filtered data are available at DRYAD doi:XXXXXX. All feces identified as common leopard with the *PantF/PantR* specific primers were further confirmed by the sequencing and common leopard sequences were observed in all the samples with an overall relative frequency of 0.16. This means that the 2µM concentration of blocking oligonucleotide used in this experiment reduced but did not prevent completely the amplification of common leopard sequence. Three feces exclusively amplified common leopard sequences, failing to provide information about any prey item. Eleven different prey taxa were identified in the remaining 57 feces. A maximum of three prey items were observed in one feces sample, two prey in seven samples, while single prey items were recoded in 49 feces. Based on the frequency of occurrence of prey items in the 57 feces, the domestic goat predominated the diet (64.9 %), followed by dog (17.5 %) and cow (12.3 %). Domestic animals (goat, dog, cow, water buffalo, horse and sheep) occurred in 54 out of 57 feces, corresponding to a frequency of occurrence of 0.95, with 5 feces containing two domestic items. Fig. 2 illustrates the relative frequencies of different prey items present in the diet of common leopard.

Discussion

Feces analysis by combining universal primers for vertebrates (plus a blocking oligonucleotide specific to predator DNA) with next generation sequencing revealed the diet of the elusive and endangered common leopard in Pakistan. In 57 feces, we were able to identify a total of 11 prey species. Our results confirmed that the common leopard is a generalist predator, consuming a wide variety of prey ranging from large, medium to small ungulates, carnivores, rodents, birds and fish.

About 95 % of the leopard prey items in ANP are domestic animals and such a high predation rate may be due to several reasons. First, the lack of natural prey in ANP provokes leopard attacks on domestic stock. Second, poor husbandry and penning of livestock make domestic animals more vulnerable to leopard attacks. Finally, seasonal factors affect livestock susceptibility to predation; the sampling of the present study was done during the summer, when most of the livestock is sent to pasture for grazing mostly unattended, or occasionally guarded by dogs. It is interesting to note that dogs are the second most common leopard prey, results possibly originating from guard dogs as well as stray dogs. We also found leopard cat, red fox and Kashmir flying squirrel in the leopard diet. These species have never been reported leopard prey before. Among the 11 taxa eaten by leopards in ANP, only one species, Kashmir flying squirrel, is categorized as vulnerable (Sheikh & Molur, 2004). This species was represented by a single sample, indicating that leopards do not seem to exert significant predation pressure on threatened species, as observed in the case of leopard cat in the same area (Shehzad et al., submitted).

The rhesus monkey (*Macaca mulatta*) is common in ANP, and is generally believed as a major part of the leopard diet there (Lodhi, 2007). However our findings do not support this perception, as we did not find any monkey sequences in the leopard feces despite that our reference database contained several exemplars. There are conflicting reports about leopard predation on primates from different parts of its range in Asia and Africa. According to the type of habitat, primates are reported to be the major prey (Cowlishaw, 1994; Isbell, 1994; Zuberbuhler & Jenny, 2002) or occasional prey of leopards (Nowell & Jackson 1996; Hayward et al., 2006). Apart from the availability, leopard selection for primates depends on several factors. First, arboreal primates are reported to be at greater risk of predation when they are more exposed (Isbell, 1994) as in forest edges or in relatively open canopy forests (Nowell & Jackson, 1996), because such landscape structures force primates to travel on the ground, which

makes them more susceptible to predation by leopards. Second, inter-specific social dominance determines rights in food and space, and tigers are known to socially dominate over leopards (Seidensticker, 1976). In areas where both species co-exist, tigers attack more ungulate prey, and consequently leopards are left to eat higher proportion of primates (Schaller, 1967; Seidensticker, 1983). Third, Seidensticker (1983) argues that intensity of predation on primates correlates with the availability and abundance of alternate prey. For example, in Royal Chitwan National Park, Nepal, predators (tiger and leopard) kill primates only occasionally because of the abundant and diverse alternate prey base that is available (Seidensticker, 1976). Finally, leopards killed primates more often in areas where domestic livestock sources were not available (like in Meru-Betiri, Indonesia, Seidensticker & Suyono, 1980) as compared to areas where livestock is abundant (Royal Chitwan National Park, Seidensticker, 1976). The close canopy of temperate forest in ANP, which makes capturing monkeys difficult, and the availability of alternate abundant prey in the form of livestock are likely the factors accounting for the absence of monkeys in the 57 feces of leopard analyzed.

It is clear from the present findings that the leopard population in ANP sustains on domestic animals, and in fact actual loss of livestock may be even higher than that observed in the diet since leopards are considered to be surplus killers and have been reported to kill up to 22 sheep in single attacks (Sangay & Vernes, 2008). This high level of predation has adverse effects on the small local economy and sparks serious conflicts between the leopard and local community. Mitigation of conflict between the leopard and local community is of the utmost importance. Failure to address this issue will jeopardize leopard conservation in ANP.

A two-step strategy should be adopted to motivate the local community to contribute to efforts for common leopard conservation. Immediate measures may include the introduction of incentives that should lead to a better acceptance of the predator. These incentives could be support for livestock insurance, for vaccination programs, and for improving livestock guarding facilities. Mishra et al. (2003) have shown that incentive programs can potentially strengthen conservation efforts by changing people's attitudes towards conservation. One such program is insurance for livestock in affected communities, an initiative that has been tested in snow leopard conservation programs. Mishra et al. (2003) highlighted the significance of insurance programs to protect the snow leopard and its prey in the Indian Himalaya and Mongolia. As a part of the

program, livestock holders protect snow leopards and leave some space for its prey to graze.

Livestock losses due to disease are reported to be substantially higher than predation by carnivores (Uphyrkina et al., 2002; Dar et al., 2009) and impose a significant economic burden on rural communities. In Machiara National Park, Pakistan, the economic value of livestock losses was estimated to be US$ 137 183; 72.7 % of that loss was due to disease caused mortalities (Dar et al., 2009). Protecting livestock from disease can reduce overall mortality rates and economic impacts on the community. The livestock vaccination program has been used as an effective conservation tool to protect snow leopards in northern Pakistan (Nawaz, 2009). In this program, the community is trained in animal health and husbandry practices and are provided vaccines against prevailing diseases at subsidized rates. In turn, the community agrees to tolerate predation losses and protect wildlife. As a joint venture between the government, community and the Snow Leopard Trust, this program is being implemented in 12 communities in northern Pakistan (Nawaz, 2009), and could be replicated in ANP. The ANP is one of the major tourism nodes in Pakistan; millions of people visit it every year to enjoy its natural beauty and wilderness. The park entry fee and other revenues from associated tourism could help lessen the heat of conflict if shared with the local community.

Measures to curtail predation incidences can help reduce conflicts and make conservation efforts more efficient and cost-effective if they go hand in hand with the incentive programs. For example, improvements in penning and guarding facilities can bring significant reduction in predation losses in livestock enclosures. Moreover, increasing vigilance at night and guarding livestock during grazing can reduce the risk of leopard attacks.

Awareness and education programs catering to the local community, students, and park visitors can promote tolerance for the leopard and build support for conservation efforts. Mass media can be used to disseminate the information about conservation issues.

In a second step, attention should be paid to the re-introduction and protection of wild prey in ANP. Depletion of natural prey due to poaching is the main reason for the leopard's dependence on livestock for food in ANP, and thus the major source of conflict in the area. Re-introduction of wild prey species like musk deer, barking deer and goral may help resolve these issues in the longer run. However such efforts would

only be successful if poaching of these species were checked in ANP and its surrounding areas through strict law enforcement.

Conclusions

The leopard is thought to be a major predator of livestock in ANP, and its frequency of killing domestic animals has increased since the creation of park. Livestock losses by predation have an adverse effect on the small economy of the local community as well as on leopard conservation efforts. Mitigation of conflicts between the leopard population and the local community is of the utmost importance for leopard conservation in the area. Due to the lack of scientific information, the real assessment of the conflict between leopard population and local community in ANP is not possible. The findings of the present study have clearly demonstrated that at least in the summer, the leopard population in ANP is largely reliant upon domestic animals for its survival. A two-step strategy should be adopted to mitigate the conflict, first by introducing incentives in and around the protected area, and then by making wild prey available again. We hope that the results of this study will rapidly stimulate the implementation of such a conservation strategy and ensure the survival of the leopard population in Pakistan.

Acknowledgements

The authors are grateful to Alice Valentini, Christian Miquel, Carole Poillot, Delphine Rioux and Ludovic Gielly for their valuable comments and suggestions throughout the course of this study. Wasim Shehzad and Tiayyba Riaz were supported by PhD scholarship from HEC (Higher Education Commission), Government of Pakistan. The Khyber-Pakhtunkhwa Wildlife Department, Himalayan Wildlife Foundation and Snow Leopard Trust organized sample collections in Ayubia National Park.

Conflict of interest statement

Eric Coissac, Tiayyba Riaz, and Pierre Taberlet are coinventors of a pending French patent on the primer pair named 12SV5F and 12SV5R and on the use of the amplified fragment for identifying vertebrate species from environmental samples. This patent only restricts commercial applications and have no impact on the use of this method by academic researchers.

References

AL-JOHANY, A.M.H. (2007) Distribution and conservation of the Arabian Leopard Panthera pardus nimr in Saudi Arabia. *Journal of Arid Environments,* 68, 20-30.

BAGCHI, S. & MISHRA, C. (2006) Living with large carnivores: predation on livestock by the snow leopard (*Uncia uncia*). *Journal of Zoology,* 268, 217-224.

BAILEY, T.N. (2005) *The African Leopard: Ecology and Behaviour of a Solitary Felid,* Blackburn Press, New Jersey.

BELLEMAIN, E., CARLSEN, T., BROCHMANN, C., COISSAC, E., TABERLET, P. & KAUSERUD, H. (2010) ITS as an environmental DNA barcode for fungi: an *in silico* approach reveals potential PCR biases. *BMC Microbiology,* 10, 9.

BLAXTER, M.L. (2004) The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London Series B,* 359, 669-679.

BROWNSTEIN, M.J., CARPTEN, J.D. & SMITH, J.R. (1996) Modulation of non-templated nucleotide addition by *Taq* polymerase: primer modification that facilitate genotyping. *BioTechniques,* 20, 1004-1010.

CONOVER, M.R., PITT, W.C., KESSLER, K.K., DUBOW, T.J. & SANBORN, W.A. (1995) Review of human injuries, illnesses, and economic-losses caused by wildlife in the United-States. *Wildlife Society Bulletin,* 23, 407-414.

COWLISHAW, G. (1994) Vulnerability to predation in baboon populations. *Behaviour,* 131, 293-304.

DAR, N.I., MINHAS, R.A., ZAMAN, Q. & LINKIE, M. (2009) Predicting the patterns, perceptions and causes of human-carnivore conflict in and around Machiara National Park, Pakistan. *Biological Conservation,* 142, 2076-2082.

DISTEFANO, E. (2005) *Human-Wildlife conflict worldwide: a collection of case studies, analysis of management strategies and good practices*. Food and Agricultural Organization of the United Nations (FAO), Sustainable Agriculture and Rural Development (SARD) paper (web document) http://www.fao.org/sard/common/ecg/1357/en/hwc_final.pdf.

FAROOQUE, M. (2007) *Management plan of Ayubia National Park,* Khyber Pakhtunkhwa Wildlife Department, Peshawar.

FICETOLA, G., COISSAC, E., ZUNDEL, S., RIAZ, T., SHEHZAD, W., BESSIERE, J., TABERLET, P. & POMPANON, F. (2010) An In silico approach for the evaluation of DNA barcodes. *BMC Genomics,* 11, 434.

GRAHAM, K., BECKERMAN, A.P. & THIRGOOD, S. (2005) Human-predator-prey conflicts:

ecological correlates, prey losses and patterns of management. *Biological Conservation,* 122, 159-171.

HAYWARD, M.W., HENSCHEL, P., O'BRIEN, J., HOFMEYR, M., BALME, G. & KERLEY, G.I.H. (2006) Prey preferences of the leopard (*Panthera pardus*). *Journal of Zoology,* 270**,** 298-313.

INSKIP, C. & ZIMMERMANN, A. (2009) Review Human-felid conflict: a review of patterns and priorities worldwide. *Oryx,* 43, 18-34.

ISBELL, L.A. (1994) Predation on primates: Ecological patterns and evolutionary consequences. *Evolutionary Anthropology,* 3, 61-71.

KABIR, M. & WASEEM, M. (2010) *Human Leopard Conflict Assessment In Pir Lasora National Park Kotli, Azad Jammu and Kashmir, Pakistan*. SGP WWF- Pakistan.

LODHI, A. (2007) *Conservation of leopards in Ayubia National Park, Pakistan*. MS thesis. The University of Montana.

MADHUSUDAN, M.D. (2003) Living amidst large wildlife: Livestock and crop depredation by large mammals in the interior villages of Bhadra Tiger Reserve, south India. *Environmental Management,* 31, 466-475.

MAGNUSON, V.L., ALLY, D.S., NYLUND, S.J., KARANJAWALA, Z.E., RAYMAN, J.B., KNAPP, J.I., LOWE, A.L., GHOSH, S. & COLLINS, F.S. (1996) Substrate nucleotide-determinated non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. *BioTechniques,* 21, 700-709.

MAIKHURI, R.K., NAUTIYAL, S., RAO, K.S. & SAXENA, K.G. (2001) Conservation policy-people conflicts: a case study from Nanda Devi biodiversity reserve (a World Heritage Site, India). *Forest Policy Economics,* 2, 355–365.

MCCARTHY, T. (2000) *Ecology and conservation of snow leopards, Gobi Brown bears and wild Bactrian camels in Mongolia*. Ph.D. thesis. University of Massachusett, Amherst.

MISHRA, C. (1997) Livestock depredation by large carnivores in the Indian trans-Himalaya: conflict perceptions and conservation prospects. *Environmental Conservation,* 24, 338-343.

MISHRA, C., ALLEN, P., MCCARTHY, T., MADHUSUDAN, M.D., BAYARJARGAL, A. & PRINS, H.H.T. (2003) The role of incentive programs in conserving the snow leopard. *Conservation Biology,* 17, 1512-1520.

MISHRA, C., VAN WIEREN, S.E., KETNER, P., HEITKONIG, I.M.A. & PRINS, H.H.T. (2004) Competition between domestic livestock and wild bharal *Pseudois nayaur* in the

Indian Trans-Himalaya. *Journal of Applied Ecology,* 41, 344-354.

NAWAZ, M.A. (2009) *Snow leopard-friendly program set to expand in Pakistan*. Newsletter of SLT, Washington.

NEEDLEMAN, S.B. & WUNSCH, C.D. (1970) A general method applicable to search for similarities in amino acid sequence of two proteins. *Journal of Molecular Biology,* 48, 443-453.

NOWELL, K. & JACKSON, P. (1996) Wild cats, status survey and conservation action plan. IUCN Publications, Burlington Press, Cambridge.

OGRA, M.V. (2008) Human-wildlife conflict and gender in protected area borderlands: A case study of costs, perceptions, and vulnerabilities from Uttarakhand (Uttaranchal), India. *Geoforum,* 39, 1408-1422.

OGRA, M. & BADOLA, R. (2008) Compensating Human-Wildlife Conflict in Protected Area Communities: Ground-Level Perspectives from Uttarakhand, India. *Human Ecology,* 36, 717-729.

OLI, M.K., TAYLOR, I.R. & ROGERS, M.E. (1994) Snow leopard *Panthera uncia* predation of livestock: An assessment of local perceptions in the Annapurna Conservation Area, Nepal. *Biological Conservation,* 68, 63-68.

PACKER, C., IKANDA, D., KISSUI, B. & KUSHNIR, H. (2005) Lion attacks on humans in Tanzania - Understanding the timing and distribution of attacks on rural communities will help to prevent them. *Nature,* 436, 927-928.

PEGARD, A., MIQUEL, C., VALENTINI, A., COISSAC, E., BOUVIER, F., FRANCOIS, D., TABERLET, P., ENGEL, E. & POMPANON, F. (2009) Universal DNA-Based Methods for Assessing the Diet of Grazing Livestock and Wildlife from Feces. *Journal of Agricultural and Food Chemistry,* 57, 5700-5706.

PRATER, S.H. (1993) *The book of Indian animal,* Oxford University Press.

RIAZ, T., SHEHZAD, W., VIARI, A., POMPANON, F.O., TABERLET, P. & COISSAC, E. (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, doi:10.1093/nar/gkr732.

ROBERTS, T.J. (1997). Mammals of Pakistan. Oxford University Press, Karachi.

SABERWAL, V.K., GIBBS, J.P., CHELLAM, R. & JOHNSINGH, A.J.T. (1994) Lion human conflict in the Gir Forest, India. *Conservation Biology,* 8, 501-507.

SANGAY, T. & VERNES, K. (2008) Human-wildlife conflict in the Kingdom of Bhutan: Patterns of livestock predation by large mammalian carnivores. *Biological Conservation,* 141, 1272-1282.

SCHALLER, G.B. (1967) The deer and the tiger, a study of wildlife in India. (ed U.o.C. Press), University of Chicago Press, Chicago.

SEIDENSTICKER, J. (1976) On the ecological separation between tigers and leopards. *Biotropica,* 8, 225-234.

SEIDENSTICKER, J. (1983) Predation byPanthera cats and measures of human influence in habitats of South Asian monkeys *International Journal of Primatology,* 4, 323-326.

SEIDENTSTICKER, J. & SUYONO, I. (1980) *The Javan Tiger and the Meru-Betiri Reserve, a Plan for Management*. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

SEKHAR, N.U. (1998) Crop and livestock depredation caused by wild animals in protected areas: the case of Sariska Tiger Reserve, Rajasthan, India. *Environmental Conservation,* 25, 160-171.

SHEHZAD, W., RIAZ, T., NAWAZ, M.A., MIQUEL, C., POILLOT, C., SHAH, S.A., POMPANON, F., COISSAC, E. & TABERLET, P. (2011) A universal approach for carnivore diet analysis based next generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology,* submitted.

SHEIKH, K.M. & MOLUR, S. (eds.) (2004) *Status and Red List of Pakistan's Mammals. Based on the Conservation Assessment and Management Plan Workshop,* IUCN Pakistan.

STUDSROD, J.E. & WEGGE, P. (1995) Park-people relationships - the case of damage caused by park animals around the Royal-Bardia-National-Park, Nepal. *Environmental Conservation,* 22, 133-142.

TREVES, A. (2009) Hunting for large carnivore conservation. *Journal of Applied Ecology,* 46, 1350-1356.

TREVES, A., WALLACE, R.B. & WHITE, S. (2009) Participatory Planning of Interventions to Mitigate Human-Wildlife Conflicts. *Conservation Biology,* 23, 1577-1587.

UPHYRKINA, O. & O'BRIEN, S.J. (2002) Applying molecular genetic tools to the conservation and action plan for the critically endangered Far Eastern leopard (*Panthera pardus orientalis*). In *International Congress on Biodiversity Conservation and Management* pp. S93-S97. Editions Scientifiques Medicales Elsevier, Vouziers, France.

VALENTINI, A., MIQUEL, C., NAWAZ, M.A., BELLEMAIN, E., COISSAC, E., POMPANON, F., GIELLY, L., CRUAUD, C., NASCETTI, G., WINCKER, P., SWENSON, J.E. & TABERLET,

P. (2009) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trn*L approach. *Molecular Ecology Resources,* 9, 51-60.

VESTHEIM, H. & JARMAN, S.N. (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples - a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology,* 5, 11.

WANG, S.W. & MACDONALD, D.W. (2006) Livestock predation by carnivores in Jigme Singye Wangchuck National Park, Bhutan. *Biological Conservation,* 129, 558-565.

WHITE, L.A. & GEHRT, S.D. (2009) Coyote Attacks on Humans in the United States and Canada. *Human Dimensions of Wildlife: An International Journal,* 14, 419-432.

ZUBERBUHLER, K. & JENNY, D. (2002) Leopard predation and primate evolution. *Journal of Human Evolution,* 43, 873-886.

Figures Legends

Fig. 1 Sampling locations of common leopard cat feces in Ayubia National Park (ANP).

Fig. 2 Composition of the various prey items consumed and their relative frequency in the diet of common leopard at Ayubia National Park.

Tables

Table 1 Sequences of the oligonucleotides used in the study. *PantF*, *PantR*, *12SV5F*, and *12SV5R* are amplification primers. *PantB* is a blocking oligonucleotide. The length of amplified fragments (excluding length of primers) with *Pant* and *12SV5* primer pairs were 79 bp and ≥ 101 bp respectively

Name	Oligonucleotide sequence (5'-3')	Reference
PantF	GTCATACGATTAACCCGG	Ficetola et al., 2010
PantR	TGCCATATTTTTATATTAACTGC	Ficetola et al., 2010
12SV5F	TTAGATACCCCACTATGC	Riaz et al., 2011
12SV5R	TAGAACAGGCTCCTCTAG	Riaz et al., 2011
PantB	CTATGCTTAGCCCTAAACCTAGATAGTTA GCCCAAACAAAACTAT-C3	This study

Fig. 1

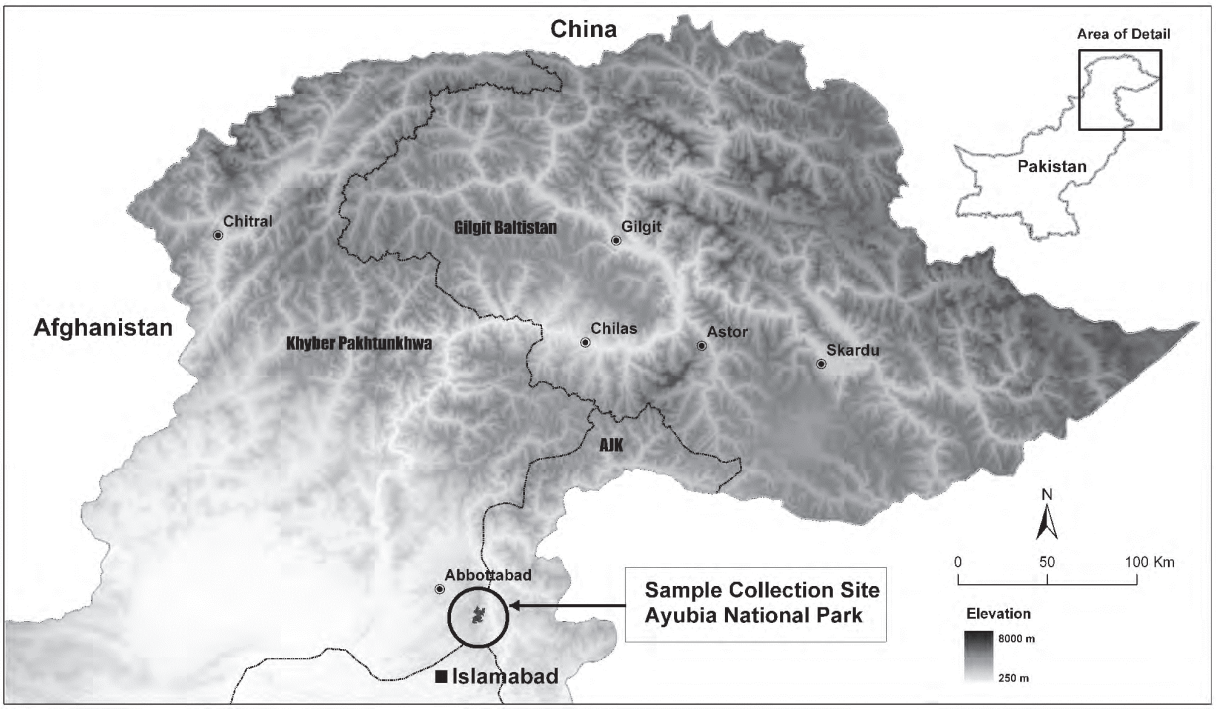
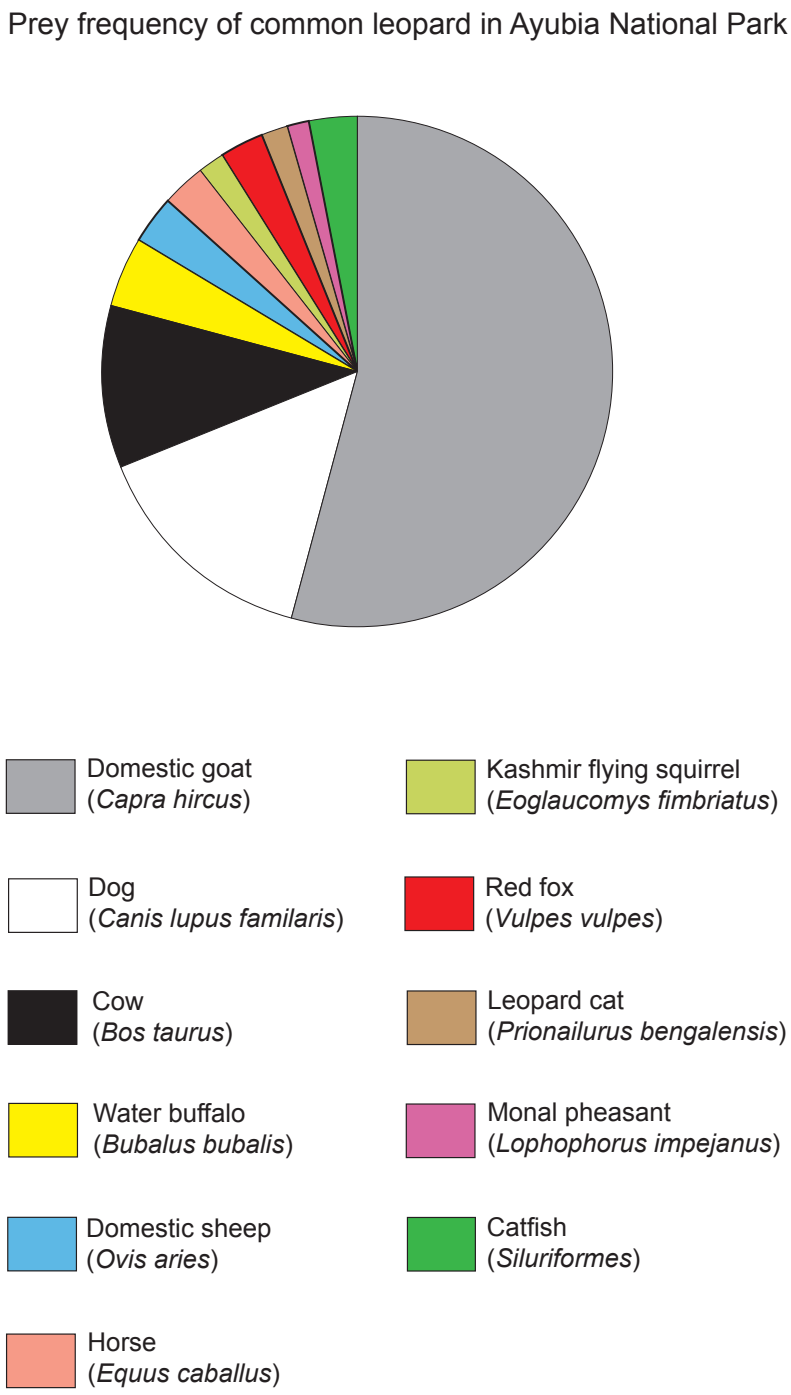

China
Area of Detail
Pakistan
Chitral
Gilgit Baltistan
Gilgit
Afghanistan
Khyber Pakhtunkhwa
Chilas
Astor
Skardu
AJK
N
0
50
100 Km
Abbottabad
Sample Collection Site
Ayubia National Park
Elevation
8000 m
250 m
Islamabad

Fig. 2


Prey frequency of common leopard in Ayubia National Park
Domestic goat (Capra hircus)
Kashmir flying squirrel (Eoglaucomys fimbriatus)
Dog (Canis lupus familaris)
Red fox (Vulpes vulpes)
Cow (Bos taurus)
Leopard cat (Prionailurus bengalensis)
Water buffalo (Bubalus bubalis)
Monal pheasant (Lophophorus impejanus)
Domestic sheep (Ovis aries)
Catfish (Siluriformes)
Horse (Equus caballus)

5. Prey preference of snow leopard (*Panthera uncia*) in South Gobi, Mongolia

WASIM SHEHZAD[1], THOMAS M. McCARTHY[2], FRANÇOIS POMPANON[1], LKHAGVAJAV PUREVJAV[3], ERIC COISSAC[1], TIAYYBA RIAZ[1], PIERRE TABERLET[1*]

1 Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 9, France

2 Panthera, 8 West 40th Street, 18th Floor, New York, New York 10018, USA

3 Snow Leopard Conservation Fund, Post Office 44, PO Box 774, Ulaanbaatar 44, Mongolia

Received:

Keywords: blocking oligonucleotide, diet analysis, DNA barcoding, next generation sequencing, *Panthera uncia*.

Corresponding author: Pierre Taberlet, Laboratoire d'Ecologie Alpine, CNRS-UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 09, France; Tel: +33(0)4 76 63 54 33; Fax: +33(0)4 76 51 42 79; E-mail: pierre.taberlet@ujf-grenoble.fr

Abstract

Accurate information about the diet of large carnivores that are elusive and inhabit inaccessible terrain, is required to properly design conservation strategies. Predation on livestock and retaliatory killing of predators have become serious issues throughout the range of the snow leopard. Several feeding ecology studies of snow leopards have been conducted using classical approaches. These techniques have inherent limitations in their ability to properly identify both snow leopard feces and prey taxa. To examine the frequency of livestock prey and nearly-threatened argali in the diet of the snow leopard, we employed the recently developed DNA-based diet approach to study a snow leopard population located in the Tost Mountains, South Gobi, Mongolia. After DNA was extracted from the feces, a region of ~100 bp long from mitochondrial 12S rRNA gene was amplified, making use of universal primers for vertebrates and a blocking oligonucleotide specific to snow leopard DNA. The amplicons were then sequenced using a next-generation sequencing platform. We observed a total of five different prey items from 81 fecal samples. Siberian ibex predominated the diet (in 70.4 % of the feces), followed by domestic goat (17.3 %) and argali sheep (8.6 %). The major part of the diet was comprised of large ungulates (in 98.8 % of the feces) including wild ungulates (79 %) and domestic livestock (19.7 %). The findings of the present study will help to understand the feeding ecology of the snow leopard, as well as to address the conservation and management issues pertaining to this wild cat.

Introduction

Apex predators play a key role in maintaining biodiversity in an ecosystem, through population dynamics and trophic cascades [1,2]. Despite a relatively wide distribution across 12 central Asian countries, information about snow leopards is scarce due to their remote habitat and cryptic nature [3]. The International Union for the Conservation of Nature (IUCN) Red List of Threatened Species has identified the snow leopard as being endangered since 1988; Appendix I of the Convention on International Trade in Endangered Species (CITES) has listed it as such since 1975. Its population is estimated to be between 4,500 and 7,500 throughout its range [4], with 800 to 1,700 in Mongolia [5]. Snow leopard populations are declining across most of its range. Primary threats include habitat degradation and fragmentation due to livestock grazing and human population expansion, poaching for pelts and bones, killings of snow leopards in retribution for predation on livestock, and reduction of natural prey populations due to illegal and legal hunting, as well as competition from livestock [6-8].

Diet analysis helps to reveal the plasticity of a predator's ability to both use and conserve resources available to it. This knowledge is important, especially when an animal is as endangered and secretive in nature as the snow leopard [9]. Snow leopards are considered to be opportunistic predators that exploit a wide range of prey species [10]. Large ungulates (notably blue sheep, markhor, urial, ibex, goats and sheep) often represent the major constituents of the snow leopard's diet. Additional prey items that have been observed include unidentified birds and a wide variety of medium and small mammals, such as marmots and other rodents [9,11-14]. In Mongolia, Siberian ibex and argali are the natural prey of the snow leopard [15].

Detailed and accurate data on predators as well as their prey is required to assess the real magnitude of predation. This knowledge is imperative for the design of balanced conservation strategies.

To date, the diet of the snow leopard has been analyzed using different classical methods. Inference from field surveys, questionnaires and interviews with members of the local community can give an assessment of snow leopard predation and has been effectively used in some studies [e.g. 16,17]. But such studies may only represent public opinion if they lack scientific confirmation. Radio telemetry provides a realistic opportunity to study animal movements, home range, pattern of habitat utilization, social

organization [5] and to document predation by snow leopards [9,18]. Although locating the remains of killed prey in high, steep terrain is extremely difficult [19], current GPS collaring studies of snow leopards in Mongolia (Panthera/Snow Leopard Trust unpublished data) have successfully identified kill sites and prey remains in over 250 instances between 2008 and 2011. However, that study is unique in its success and is not easily replicated across the snow leopard's broad range. Additionally, diet information from such studies remains site specific.

Fecal examination may then represent the most readily available and easily collected source of diet information [20]; this technique has been previously used to study snow leopard diet [7,9,11,14]. Such diet analysis requires the identification of undigested remains, bones, teeth or hair in feces. There have been two potential problems relating to feces examination of the snow leopard. The first relates to the accurate identification of snow leopard feces in the field, while the second deals with the limitations of accurately identifying the prey taxa. Large bones and teeth are generally fragmented and therefore difficult to identify [21]. Hair examination is commonly done through comparisons of reference specimens with salvaged hair mounts [21]. However, this method is laborious and time consuming. Hairs from the same animal may also vary in structure according to their location within its fur. Similarly, hair from several related species may possess similar characteristics [21]. Finally, the lack of reference specimens can prohibit accurate diagnosis.

Snow leopard feces are identified in the field mostly on the basis of color, shape, location, pugmarks, scrapes, or the remains of prey species near the feces [11,14]. However, carnivore feces are quite similar in their morphological characteristics; it is therefore not always easy to differentiate the feces among the sympatric carnivores [14, 22, 23]. To address this problem, some studies [e.g. 24,25] have used detection dogs trained to locate the feces of endangered species. In another study, scat detection dogs are being trained to distinguish snow leopard feces from other non-target feces after they have been collected in the field and shipped to the USA, thus eliminating the cost and complications of bringing a dog to the field. This may prove to be cost-effective for studies of snow leopards in which finding feces in the field is not difficult, but making a correct species identification is critical (McCarthy & Parker, unpublished data). But this approach is also prone to sample misidentification. Assumptions based on such errors

may lead to incorrect conclusions regarding the conservation implications for these endangered species. Alternatively, feces collection in the field has been validated by genetic analysis [26,27]. So far, two studies [7,9] have identified snow leopard feces genetically and then determined prey consumed using classical approaches.

Livestock depredation has become a real challenge in central Asia throughout the range of the snow leopard [28]. Snow leopards are thought to be one of the major killers of livestock, which results in hostility towards the animal from local communities [16,29,30] and retribution killings of snow leopards [14,31]. In a survey conducted in the four regions of Mongolia, 14 % of livestock holders have admitted to hunting snow leopard to carry out retribution killings of their livestock [32]. Namgail et al. [17] have stated that 38 % of the total livestock losses in Ladakh, India, can be attributed to snow leopards. Similarly, high proportions (40 - 58 %) of livestock in snow leopard diet have been reported in two regions of India [14]. Due to these factors, local support of conservation efforts for this wild cat is questionable across its range.

This study aims to test two hypotheses about the snow leopard population from the Tost Mountains, South Gobi, Mongolia (i) Does livestock constitute the major part of the snow leopard diet? (ii) Does the snow leopard predate on threatened argali sheep? In our effort to address the potential limitations of the methods available to analyze the diet of snow leopards, we have benefited from the recent development of a DNA-based universal approach for diet analysis [33], which combines universal primers for vertebrates with a blocking oligonucleotide specific to snow leopard DNA and subsequent high-throughput next-generation sequencing. The results of the present study precisely identify all the prey items consumed, thus providing important knowledge for the conservation and management of this wild cat.

Results

Of 203 putative snow leopard fecal samples collected in the field of Tost Mountains South Gobi Mongolia, 88 (43.3 %) samples were identified as originating from this species using snow leopard specific primers (*UnciF/UnciR*); these were selected for further experimentation. After assembling the forward and reverse reads, and filtering for primers and tags, we obtained a total of 1900638 sequences for these 88 feces samples, corresponding to 173770 unique sequences (DRYAD entry: doi:XXXXXX). Removing sequences shorter than 60 bp and with a total count inferior to 100 reduced the dataset to 463 sequences. Also removed were 364 sequences that were never "head" (most common sequence among all sequences that can be linked with a single indel or substitution) or "singleton" (no other variant with a single difference in the relevant PCR product), as were 74 sequences with a total count among all samples inferior to 1000 (these sequences only corresponded to variants of more common sequences). This threshold has been determined in order to avoid duplicative species identifications in the results list and with a lower identity for sequences below the threshold. Finally, 19 sequences were removed that lacked perfect identification from the reference database or were that of prey that had already been identified with higher count, leaving six (including one *Panthera uncia* and five prey sequences) MOTU [Molecular Operational Taxonomic Unit; 34]. Table 1 presents the overview of sequence counts at different stages of the analysis. The filtered data are available at DRYAD doi:XXXXXX. All feces identified as snow leopard with the snow leopard specific primers (*UnciF/UnciR*) were further confirmed by sequencing; snow leopard sequences were observed in all of the samples with an overall relative frequency of 0.26. This means that the 2µM concentration of blocking oligonucleotide used in this experiment reduced, but did not completely block the amplification of the snow leopard sequence (see Table 2 for a comparison of the results with and without blocking oligonucleotide). We observed seven samples that produced only the snow leopard DNA fragment but no prey count. Nine samples contained a sequence that corresponds to both argali and domestic sheep. These two species were prevalent in our study area. Further sequencing experiments with primer pair *OvisF/OvisR*, amplifying a 82 bp of the cytochrome *b* gene, had successfully identified argali and domestic sheep in seven and two samples, respectively.

Diet composition of snow leopard

The diet composition of snow leopards in the Tost Mountains of Mongolia was not highly diverse; a total of only five different prey items were identified in the diet. All prey taxa were identified to the species level. We observed one prey item per sample in all 81 samples, although we could not amplify any prey sequence in seven samples.

On the basis of its occurrence in feces, the Siberian ibex was observed to be the most frequent prey (70.4 %), followed by domestic goat (17.3 %) and argali sheep (8.6 %). Overall, ungulates comprised the dominant part of the diet (in 98.8 % of the feces) including wild species (79 %) and domestic livestock (19.7 %). Only one species of bird other than ungulates was recorded; a chukar partridge was recorded in one (1.2 %) fecal sample (Figure 2). Table 3 presents an overview of the snow leopard diet in the Tost Mountains, South Gobi, Mongolia, in comparison with other studies.

Discussion

Our results are in general agreement with previous studies, indicating that large ungulates (wild and domestic) represent the major part of snow leopard diet (in 38.7 - 98.8 %, of the feces). Moreover, Siberian ibex are most abundantly observed (9.1 - 70.4 %) in the diet in many parts of snow leopard range except in the regions where ibex are rare [11,12]. However, our results differ from other studies in that we are reporting high predation on wild ungulates (79 %), which is considerably higher that previous reports (which range from 12.9 - 56.9 %). It seems that snow leopard predation and plasticity depends upon the availability of its natural prey. Our study area appears to host an abundant population of wild ungulates, including Siberian ibex and argali sheep. It may be easier for snow leopard to attack wild prey than domestic, because the latter is often guarded by humans. In previously reported studies, medium and small mammals provide an important dietary supplement (3.9 - 53.3%). The absence of any medium or small mammals in our results cannot be considered to be an indication of preference, since we have no data on prey availability. Large wild cats tend to prefer large-sized prey [35-37]. Bird species represent an important element of snow leopard diet; although they are observed in low frequencies (1.2 – 15.9 %), birds are consistent in all documented studies (Table 3).

Our findings support previous studies that snow leopard tend to focus primarily on a single prey item [11,12]. We also observed only a single prey species per fecal sample. The average body weight of an adult snow leopard is about 40 - 45 kg [21,38], for which the required daily prey biomass is estimated to be 1.5 to 2.5 kg [12]. Predation on large prey is therefore sufficient to fulfill its bodily requirements for several days. McCarthy [5] estimated that snow leopards kill a large prey item every 10 - 15 days and feed on it for an average of 3 - 4 days and sometimes up to one week.

Although this study only examined vertebrate prey items consumed by snow leopards, some studies [9,11-14] have documented plant material in snow leopard feces (see also Table 3). Why snow leopards consume plant matter has not been definitively determined, although some studies [11,13,14] have suggested this phenomenon is the result of accidental ingestion while feeding on prey. The possibility has not been ruled out that plant matter fills some specific dietary need such as providing minerals or vitamins not readily gained from animal matter.

We observed seven samples that contained snow leopard DNA exclusively; no prey DNA could be amplified from them. A plausible explanation for this is found in that snow leopards may go several days between meals [5,19] and in the latter part of this interval its feces would likely contain mostly hair (from grooming) and the cat's own metabolic waste products. Another difficulty arose in differentiating between argali and domestic sheep. These two species have the same sequence when using the universal primers for vertebrates *12SV5F/12V5R*. Both these prey species are found in the study area. To remedy this, we designed a new primer pair *OvisF/OvisR*, targeting a part of the cytochrome *b* gene from mitochondrial DNA. It showed consistent variation between these two potential prey species and helped to discern argali and domestic sheep remains.

DNA-based techniques provide a powerful means to study the feeding ecology of wild and cryptic species like the snow leopard. By using universal primers for vertebrates, limiting snow leopard sequences with a blocking oligonucleotide, and then using next generation sequencing, we were able to precisely identify all prey items to the species level. Traditional methods of snow leopard diet analyses via fecal samples have been unable to identify soft and well-digested components or the remains of specific bird species. Our approach has an obvious advantage in that virtually no vertebrate prey remained unidentified as the *12SV5F/12V5R* primers amplify 98% of all vertebrates [47].

Conservation implications for snow leopards in Mongolia

A clear and unambiguous understanding of an endangered carnivore's diet is crucial for conservation planning for the species. To date, despite numerous studies, the diet of the snow leopard has been inadequately assessed. This has been true for at least two reasons: misidentifying feces as being those of snow leopards, and the inherent inaccuracies of classic macro- or micro-histological examination of fecal content. This study shows that both potential shortcomings can now be overcome through genetic techniques.

A better understanding of the diet of the snow leopard will allow us to more accurately assess the level of conflict between the cats and pastoralists who rightly or wrongly attribute their livestock depredation losses to snow leopards. Mitigating measures can then be designed that address a real, as opposed to a perceived, conflict.

In this study, argali, the largest wild mountain sheep, represented 8.6 % of diet of snow leopards in the Tost Mountains. Argali are listed as "nearly threatened" in the

IUCN Red List. Threats to the species, other then predation by snow leopards, include loss or degradation of habitat due to competition with domestic sheep and illegal hunting for meat and horns [39,40]. In consideration of their prevalence in snow leopard diet, adequate conservation strategies are required to protect and increase the existing populations of argali in the Tost Mountains and elsewhere in the snow leopard range [41-43]. Illegal hunting should strictly be discouraged while habitat restoration for this wild sheep should be pursued.

The techniques we describe here can also be employed to help assess opportunities to increase snow leopard numbers in areas where they have been reduced. Knowledge of diet composition and prey availability in such instances would help conservationists determine if adequate wild prey is available to support hoped-for increases in snow leopard populations. This would help avoid situations where increasing leopard numbers only result in escalating conflicts with livestock and humans, dooming the effort to failure. Conversely, where conflict is already high and conservation efforts focus on reducing livestock depredation (predator-proof corrals, better guard dogs, etc.), an accurate assessment of current diet composition and wild prey availability would help avert unintended stress to snow leopards already facing inadequate food supplies to sustain their existing numbers.

Materials and methods

Study area and sample collection

During the summer of 2009, 203 fecal samples were collected in the Tost Mountains of South Gobi province, Mongolia (Figure 1). The fecal collection, conducted by Panthera and the Snow Leopard Trust, was part of an ongoing long-term study of snow leopard ecology [44]. All necessary permits were obtained for the described field studies. Sampling authorizations in the South Gobi study area were provided by the Mongolian Ministry of Nature, Environment and Tourism. All samples were in 10 ml vials with ~ 6 ml of silica gel. Each sample was later divided into three sub-samples for separate analysis. One sub-sample was submitted to the Global Felid Genetics Program at the American Museum of Natural History (AMNH) for species and individual identification through microsatellite analyses. A second sub-sample was sent to Working Dogs for Conservation (Bozeman, Montana, USA) to help train and test scat detection dogs in species identification. A third sub-sample was sent to the Laboratoire d'Ecologie Alpine (Université Joseph Fourier, Grenoble, France) for dietary analyses.

DNA extraction

All extractions were performed in a room reserved for the extraction of degraded DNA. Total DNA was extracted from about 15 mg of feces with the DNeasy Blood and Tissue Kit (QIAgen GmbH, Hilden, Germany), following the manufacturer's instructions, with a slight modification at the beginning of the protocol as described by Shehzad et al. [33]. The DNA extracts were recovered in a total volume of 250 µL. Mock extractions without samples were systematically performed to monitor possible contaminations.

Designing primer pairs for snow leopard diet study

Identification of predator species

A primer pair *UnciF/UnciR* highly specific to the snow leopard was designed on the 12S mitochondrial rRNA gene (Table 4), with the 3'-end of each primer as different as possible from other species. The specificity of this primer pair, amplifying a 68 bp fragment, was validated *in silico* by using the *ecoPCR* program [45,46], with the following parameters: a perfect match on the two last nucleotides, and a maximum of three mismatches on the remaining nucleotides. Using these parameters, only *P. uncia* mitochondrial 12S gene was recovered. This *in silico* validation confirmed that snow

leopard specific primers (*UnciF/UnciR*) should unambiguously identify snow leopard feces. The PCRs were carried out in a total volume of 20 µl with 8 mM Tris-HCl (PH 8.3), 40 mM KCl, 2 mM $MgCl_2$, 0.2 µM of each primer, BSA (5 µg), 0.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) using 2 µL as DNA template. The PCR conditions chosen were an initial 10 min denaturation step at 95°C, followed by 45 cycles of denaturation at 95°C for 30 s and annealing at 53°C for 30 s. Thus, the primary identification of the samples was done on the basis of the presence of a PCR product of the suitable length revealed by electrophoresis on a 2 % agarose gel. The samples successfully amplified by the snow leopard specific primer pair (*UnciF/UnciR*) were selected for further analyses.

Universal primer pair for vertebrates 12SV5F/12SV5R

We used the primer pair *12SV5F*/*12SV5R* (Table 4) designed by the *ecoPrimers* program [47]. The *ecoPrimers* scans whole genomes to find new barcode markers and their associated primers. This program optimizes two quality indices that measure the taxonomical coverage and the potential discrimination power in order to select the most efficient markers, according to specific experimental constraints such as marker length or targeted taxa. This universal primer pair for vertebrates represents the best choice found by *ecoPrimers* among short barcodes, according to the available vertebrate whole-mitochondrial genomes currently available. It amplifies a ~100 bp fragment of the V5 loop of the mitochondrial 12S gene with the ability to amplify short DNA fragments such as recovered from feces, and which has a high taxonomic resolution despite its short size. Using the *ecoPCR* program [45,46], and based on the 103 release of the EMBL database, this fragment unambiguously identifies 77 % of the species and 89 % of the genera.

Blocking oligonucleotide specific to snow leopard sequence

The *UnciB* (Table 4) blocking oligonucleotide sequence specific to snow leopards was designed as suggested by Vestheim & Jarman [48]. It overlaps the amplification primer *12SV5R* by six nucleotides. This blocking oligonucleotide was used to restrict the amplification of snow leopard sequences when using the universal primers that target all vertebrates.

Primer pair to distinguish Ovis aries and O. ammon

The two closely related prey species, domestic sheep (*Ovis aries*) and argali sheep (*O. ammon*), have potentially similar sequences for the amplified fragment of 12S

ribosomal RNA gene. A special primer pair *OvisF/OvisR* (Table 4) targeting cytochrome *b* gene was designed to amplify a homologous mitochondrial DNA region (~82 bp) that shows consistent variation between *Ovis aries* and *O. ammon*.

DNA amplification for diet analysis

We performed experiments without and with the blocking oligonucleotide. All DNA amplifications were carried out in a final volume of 25 µL, using 2 µL of DNA extract as template. The amplification mixture contained 1 U of AmpliTaq Gold® DNA Polymerase (Applied Biosystems, Foster City, CA), 10 mM Tris-HCl, 50 mM KCl, 2 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.1 µM of each primer (*12SV5F*/*12SV5R*) and 2 µM for *PrioB* (*PrioB* only in the experiments with blocking oligonucleotide), and 5 µg of bovine serum albumin (BSA, Roche Diagnostic, Basel, Switzerland). The PCR mixture was denatured at 95°C for 10 min, followed by 45 cycles of 30 s at 95°C, and 30 s at 60°C; as the target sequences are ~100 bp long, the elongation step was removed to reduce the +A artifact [49,50] that might decrease the efficiency of the first step of the sequencing process (blunt-end ligation).

The universal primers for vertebrates *12SV5F* and *12SV5R* were modified by the addition of specific tags on the 5′ end to allow the assignment of sequence reads to the relevant sample [51]. All the PCR products were tagged identically on both ends. These tags were composed of CC on the 5' end followed by seven variable nucleotides that were specific to each sample. The seven variable nucleotides were designed using the *oligoTag* program (www.prabi.grenoble.fr/trac/OBITools) with at least three differences among the tags, without homopolymers longer than two, and avoiding a C on the 5' end in order to allow detection of possible deletions within the tag. All of the PCR products from the different samples were first purified using the MinElute PCR purification kit (QIAGEN GmbH). They were then titrated using capillary electrophoresis (QIAxel, QIAgen GmbH, Hilden, Germany) and finally, mixed together in equimolar concentration before the sequencing step.

DNA sequencing

Sequencing was carried out on the Illumina Genome Analyzer IIx (Illumina Inc., San Diego, CA, 92121 USA), using the Paired-End Cluster Generation Kit V4 and the Sequencing Kit V4 (Illumina Inc.) and following manufacturer's instructions. A total of 108 nucleotides were sequenced on each extremity of the DNA fragments.

Sequence analysis and taxon assignation

The sequence reads were analyzed using the *OBITools* program (www.prabi.grenoble.fr/trac/OBITools). First, the direct and reverse reads corresponding to a single molecule were aligned and merged using the *solexaPairEnd* program, taking into account data quality during the alignment and the consensus computation. Then, primers and tags were identified using the *ngsfilter* program. Only sequences with perfect matches on tags and a maximum of two errors on primers were taken into account. The amplified regions, excluding primers and tags, were kept for further analysis. Strictly identical sequences were clustered together using the *obiuniq* program, keeping the information about their distribution among samples. Sequences shorter than 60 bp, or containing ambiguous nucleotides, or with a number of occurrences lower or equal to 100 were excluded using the *obigrep* program. The *obiclean* program was then implemented to detect amplification/sequencing errors, by giving each sequence within a PCR product the status of "head" (most common sequence among all sequences that can be linked with a single indel or substitution), "singleton" (no other variant with a single difference in the relevant PCR product), or "internal" (all other sequences not being "head" or "singleton", i.e. corresponding to amplification/sequencing errors). Taxon assignation was achieved using the *ecoTag* program [52]. *EcoTag* relies on the FASTA35 program [53] to find highly similar sequences in the reference database. This database was built by extracting the relevant part of the mitochondrial 12S gene from the EMBL nucleotide library (release 107) using the *ecoPCR* program [46]. A unique taxon was assigned to each unique sequence. This unique taxon corresponds to the last common ancestor node in the NCBI taxonomic tree of all the taxids annotating the sequences of the reference database that matched against the query sequence. A final filtering was carried out by removing sequences that were never "head" or "singleton", sequences that were not identified at the family level (for removing putative chimeras), and sequences with a total count among the whole dataset of less than 750 (plus the removing of an obvious human contamination). Finally, when a more precise identification was required, automatically assigned taxonomic identifications were completed manually by combining the automatic identification with distribution data of prey in the study area.

DNA amplification and sequencing to differentiate between domestic and argali sheep

Fecal samples that amplified domestic or argali sheep sequences were re-amplified with primer pair *OvisF/OvisR* and re-sequenced, using capillary sequencing to distinguish between these two species.

PCR amplifications were conducted in a 20 µl volume with 2 mM MgCl2, 0.2 mM of each dNTP, 0.2 µM of each primer (*OvisF/OvisR*) and 0.6 unit of AmpliTaq Gold Polymerase (Applied Biosystems). After a 10 min period at 95°C for polymerase activation, 45 cycles were run with the following steps: 95° C: 30 s, 65° C: 30 s, without the elongation step. The PCR products were purified using the QIAquick PCR purification kit (Qiagen GmbH). Fifteen nanograms of purified DNA from this PCR product were used to sequence the *OvisF* and *OvisR* primers separately. Sequence reactions were performed for both DNA strands by using the ABI PRISM Dye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Carlsbad, California 90028 USA) in a 20 µl volume with 0.2 µM of each primer. Twenty-five cycles were run with the following steps: 95° C: 30 s, 65° C: 30 s, without the elongation step. Excess dye terminators were removed by spin-column purification and the products underwent electrophoreses on an ABI 3I30xl PRISM DNA sequencer (Applied Biosystems) using the POP 7 polymer.

Acknowledgements

We thank Rana Bayrakci for a critical review of the manuscript and for providing numerous helpful suggestions. We also extend our thanks to Alice Valentini, Christian Miquel, Carole Poillot, Delphine Rioux and Ludovic Gielly for their valuable comments and suggestions throughout the course of this study. Wasim Shehzad and Tiayyba Riaz are supported by the Higher Education Commission (HEC) of Pakistan.

Author contributions

This experiment was conceived by WS, PT and FP. The field sampling was arranged by TMM and LP. The universal primers for vertebrates were designed by TR and EC. The bio-informatics programs to treat the NGS files were developed by EC. The experimentation was conducted by WS. The manuscript was written by WS, PT and TMM.

Funding

This work was funded by Laboratoire d'Ecologie Alpine (LECA), UMR 5553, CNRS (Centre National de la Recherche Scientifique), France. Partial funding was provided by Snow Leopard Conservation Grant Program for 2010. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of this manuscript.

Competing interest statement

Eric Coissac, Tiayyba Riaz, and Pierre Taberlet are co-inventors of a pending French patent on the primer pair named *12SV5F* and *12SV5R* and on the use of the amplified fragment for identifying vertebrate species from environmental samples. This patent only restricts commercial applications and has no impact on the use of this method by academic researchers. This does not alter our adherence to all the PLoS ONE policies on sharing data and materials.

References

1. Sergio F, Caro T, Brown D, Clucas B, Hunter J, et al. (2008) Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. Annu Rev Ecol Evol Syst 39: 1-19.
2. Baum JK, Worm B (2009) Cascading top-down effects of changing oceanic predator abundances. J Anim Ecol 78: 699-714.
3. Jackson RM, Roe JD, Wangchuk R, Hunter DO (2006) Estimating snow leopard population abundance using photography and capture-recapture techniques. Wildlife Soc B 34: 772-781.
4. Fox JL (1994) Snow leopard conservation in the wild - a comprehensive perspective on a low density and highly fragmented population. In: Fox JL, Jizeng D, editors. Proceedings of the Seventh International Snow Leopard Symposium. Xining, Qinghhai, China. pp. 3-15.
5. McCarthy T (2000) Ecology and conservation of snow leopards, Gobi Brown bears and wild Bactrian camels in Mongolia [Ph.D. thesis]: University of Massachusett, Amherst. 134 p.
6. McCarthy TM, Chapron G (2003) Snow leopard survival strategy; McCarthy TM, Chapron G, editors. Seattle, USA.: ISLT and SLN. 105 p.
7. Lovari S, Boesi R, Minder I, Mucci N, Randi E, et al. (2009) Restoring a keystone predator may endanger a prey species in a human-altered ecosystem: the return of the snow leopard to Sagarmatha National Park. Anim Conserv 12: 559-570.
8. Jackson RM, Mishra C, McCarthy TM, Ale SB (2010) Snow leopards: conflicts and conservation. In: Macdonald D, Loveridge A, editors. The Biology and Conservation of Wild Felids. pp. 417-430.
9. Anwar M, Jackson R, Nadeem M, Janečka J, Hussain S, et al. (2011) Food habits of the snow leopard *Panthera uncia* (Schreber, 1775) in Baltistan, Northern Pakistan. Eur J Wildlife Res 57: 1077-083.
10. Fox JL (1989) A review of the status and ecology of the snow leopard (*Panthera uncia*). Washington. 40 p.
11. Oli MK, Taylor IR, Rogers ME (1993) Diet of the snow leopard (*Panthera uncia*) in the Annapurna Conservation Area, Nepal. J Zool 231: 365-370.
12. Chundawat RS, Rawat GS (1994) Food habits of the snow leopard in Ladakh. In: Fox

JL, Jizeng D, editors. International Snow Leopard Trust and Northwest Plateau Institute of Biology, Seattle: International Snow Leopard Trust. pp. 127–132.

13. Lhagvasuren B, Munkhtsog B (2000) The yak population in Mongolia and its relation with snow leopards as a prey species. In: Jianlin H, Richard C, Hanotte O, McVeigh C, Rege JEO, editors. Proceedings of the Third International Congress on Yak production in central Asian highlands. Lhasa, China: International Livestock Research Institute. pp. 69-75.
14. Bagchi S, Mishra C (2006) Living with large carnivores: predation on livestock by the snow leopard (*Uncia uncia*). J Zool 268: 217-224.
15. Reading RP, Amgalanbaatar S, Lhagvasuren L (1999) Biological assessment of Three Beauties of the Gobi National Conservation Park, Mongolia. Biodivers Conserv 8: 1115-1137.
16. Mishra C (1997) Livestock depredation by large carnivores in the Indian trans-Himalaya: conflict perceptions and conservation prospects. Environ Conserv 24: 338-343.
17. Namgail T, Fox JL, Bhatnagar YV (2007) Carnivore-caused livestock mortality in Trans-Himalaya. Environ Manage 39: 490-496.
18. Grönberg E (2011) Movement patterns of snow leopard (*Panthera uncia*) around kills based on GPS location clusters. Riddarhyttan: Swedish University of Agricultural Sciences.
19. Jackson RM (1996) Home Range, Movements and Habitat Use of Snow Leopard (*Uncia uncia*) In Nepal [Ph.D. thesis]: University of London. 233 p.
20. Putman RJ (1984) Facts from feces. Mammal Rev 14: 79-97.
21. Oli MK (1993) A key for the identification of the hair of mammals of a snow leopard (*Panthera uncia*) habitat in Nepal J Zool 231: 71-93.
22. Hansen MM, Jacobsen L (1999) Identification of mustelid species: otter (*Lutra lutra*), American mink (*Mustela vison*) and polecat (*Mustela putorius*), by analysis of DNA from faecal samples. J Zool 247: 177-181.
23. Spiering PA, Gunther MS, Wildt DE, Somers MJ, Maldonado JE (2009) Sampling error in non-invasive genetic analyses of an endangered social carnivore. Conserv Genet 10: 2005-2007.
24. Long RA, Donovan TM, Mackay P, Zielinski WJ, Buzas JS (2007) Effectiveness of

scat detection dogs for detecting forest carnivores. J Wildl Manage 71: 2007-2017.

25. Vynne C, Skalski JR, Machado RB, Groom MJ, Jacomo ATA, et al. (2011) Effectiveness of Scat-Detection Dogs in Determining Species Presence in a Tropical Savanna Landscape. Conserv Biol 25: 154-162.
26. Davison A, Birks JDS, Brookes RC, Braithwaite TC, Messenger JE (2002) On the origin of faeces: morphological versus molecular methods for surveying rare carnivores from their scats. J Zool 257: 141-143.
27. Piggott MP, Taylor AC (2003) Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. Wildl Res 30: 1-13.
28. Jackson R, Wangchuk R (2001) Linking Snow Leopard Conservation and People-Wildlife Conflict Resolution: Grassroots Measures to Protect the Endangered Snow Leopard from Herder Retribution. Endangered Species UPDATES 18: 138-141.
29. Mishra C, Madhusudan MD, Datta A (2006) Mammals of the high altitudes of western Arunachal Pradesh, eastern Himalaya: an assessment of threats and conservation needs. Oryx 40: 29-35.
30. Inskip C, Zimmermann A (2009) Review Human-felid conflict: a review of patterns and priorities worldwide. Oryx 43: 18-34.
31. Hussain S (2003) The status of the snow leopard in Pakistan and its conflict with local farmers. Oryx 37: 26-33.
32. Allen P, McCarthy TM, Bayarjargal A (2002) Conservation de la panthère des neiges (*Uncia uncia*) avec les éleveurs de Mongolie. Paris, France.
33. Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, et al. Carnivore diet analysis based next generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. Mol Ecol. In press.
34. Blaxter ML (2004) The promise of a DNA taxonomy. Philosophical Transactions of the Royal Society of London Series B 359: 669-679.
35. Karanth KU, Sunquist ME (1995) Prey selection by tiger, leopard and dhole in tropical forests. J Anim Ecol 64: 439-450.
36. Radloff FGT, Du Toit JT (2004) Large predators and their prey in a southern African savanna: a predator's size determines its prey size range. J Anim Ecol 73: 410-423.
37. Macdonald DW, Loveridge AJ, Nowell K (2010) Dramatis personae: an introduction

to the wild felids. In: Macdonald DW, Loveridge AJ, editors. Biology and conservation of wild felids. New York: Oxford University Press.

38. Jalanka HH (1989) Medetomidine-Induced and ketamine-induced immobilization of snow leopards (*Panthera uncia*) - doses, evaluation, and reversal by atipamezole. J Zoo Wildlife Med 20: 154-162.
39. Reading RP, Amgalanbaatar S, Mix H, Lhagvasuren B (1997) Argali *Ovis ammon* surveys in Mongolia's South Gobi. Oryx 31: 285-294.
40. Maroney RL (2005) Conservation of argali *Ovis ammon* in western Mongolia and the Altai-Sayan. Biol Conserv 121: 231-241.
41. Shrestha R, Wegge P (2006) Determining the composition of herbivore diets in the trans-Himalayan rangelands: A comparison of field methods. Rangeland Ecol Manag 59: 512-518.
42. Khatoon R (2010) Snow leopard (*Uncia uncia*) diet selection in the Chitral area:MSc thesis, Arid Agriculture University. 90 p.
43. Namgail T, Fox JL, Bhatnagar YV (2004) Habitat segregation between sympatric Tibetan argali Ovis ammon hodgsoni and blue sheep *Pseudois nayaur* in the Indian Trans-Himalaya. J Zool 262: 57-63.
44. McCarthy T, Murray K, Sharma K, Johansson O (2010) Preliminary results of a long-term study of snow leopards in South Gobi, Mongolia. IUCN Cat News 53: 15-19.
45. Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, et al. (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. BMC Microbiol 10: 9.
46. Ficetola G, Coissac E, Zundel S, Riaz T, Shehzad W, et al. (2010) An In silico approach for the evaluation of DNA barcodes. BMC Genomics 11: 434.
47. Riaz T, Shehzad W, Viari A, Pompanon Fo, Taberlet P, et al. (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Res. doi:10.1093/nar/gkr732.
48. Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples - a case study on prey DNA in Antarctic krill stomachs. Front Zool 5: 11.
49. Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* polymerase: primer modification that facilitate

genotyping. BioTechniques 20: 1004-1010.

50. Magnuson VL, Ally DS, Nylund SJ, Karanjawala ZE, Rayman JB, et al. (1996) Substrate nucleotide-determinated non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. BioTechniques 21: 700-709.
51. Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, et al. (2009) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. Mol Ecol Resour 9: 51-60.
52. Pegard A, Miquel C, Valentini A, Coissac E, Bouvier F, et al. (2009) Universal DNA-Based Methods for Assessing the Diet of Grazing Livestock and Wildlife from Feces. J Agr Food Chem 57: 5700-5706.
53. Needleman SB, Wunsch CD (1970) A general method applicable to search for similarities in amino acid sequence of two proteins. J Mol Biol 48: 443-453.

Figures Legends

Figure 1 Location of Tost Mountain study site in South Gobi, Mongolia.

Figure 2 Relative frequencies of various prey species present in the diet, on the basis of their occurrence in feces of snow leopards from the Tost Mountains, South Gobi, Mongolia.

Tables

Table 1 Overview of the sequence counts at different stages of the analysis

	Number of reads (% of properly assembled sequences[a])	Number of unique sequences[b]
Number of properly assembled sequences	1900638	173770
Filtering sequences length>=60bp & count>=100	963714 (50.70 %)	463
Filtering for most of the PCR/sequencing errors	841459 (42.27 %)	74
Perfectly assigned taxa	725969 (38.19 %)	6

[a] Direct and reverse sequence reads corresponding to a single DNA molecule were aligned and merged, producing what we called a "properly assembled sequence".

[b] Strictly identical sequences correspond to “unique sequence”.

Table 2 Comparison of PCR amplification without and with blocking oligonucleotide (DRYAD entry: doi:XXXXXX).

	Snow leopard sequence (%)	Other sequences (%)
Experiments without blocking oligonucleotide	87.45	12.55
Experiments with blocking oligonucleotide	25.76	74.24

Table 3 A comparison of frequency of occurrence (percent) of various prey items in the feces of snow leopard from various regions of its range

Prey consumed	Present study South Gobi, Mongolia (n=81)	Anwar *et al.* (2011) Baltistan, Pakistan (n=49)	Bagchi & Mishra (2006) Pin Valley, India (n=51)	Bagchi & Mishra (2006) Kibber, India (n=44)	Lhagvasuren & Munkhtsog (2000) Uvs & South Gobi, Mongolia (n=168)	Chundawat & Rawat (1994) Ladakh, India (n=173)	Oli *et al.* (1993) Manang, Nepal (n=213)
Wild ungulates							
Argali	8.6	-	-	-	-	-	-
Blue sheep	-	-	-	20.5	-	23.4	51.6
Goitered gazelle	-	-	-	-	3.6	-	-
Ladakh urial	-	-	-	-	-	0.4	-
Markhor	-	3.2	-	-	-	-	-
Red deer	-	-	-	-	2.4	-	-
Roe deer	-	-	-	-	0.6	-	-
Siberian ibex	70.4	9.7	56.9	9.1	38.7	-	-
Meso & small mammals							
Hare	-	-	3.9	6.8	1.2	3.1	-
Weasel	-	-	-	-	-	-	4.7
Marmots	-	-	-	-	1.2	9.8	20.7
Marten	-	-	-	-	-	-	3.8
Red fox	-	-	-	-	-	4.3	0.9
Pika	-	-	-	-	5.9	-	15.9
Rodents	-	-	-	-	0.6	-	-
Royale's vole	-	-	-	-	-	-	7.5
Domestic livestock							
Cattle & Yak	-	8.6	2.0	6.8	4.8	1.2	14.1
Donkey	-	-	3.9	13.6	-	0.4	-
Goat	17.3	11.8	3.9	9.0	3.6	10.2	0.5
Horse	-	-	11.8	4.5	5.4	0.8	2.8
Sheep	2.5	16.1	2.0	4.5	17.3	2.3	0.5
Birds	1.2	2.2	-	15.9	2.4	3.1	1.4
Insects	-	-	-	-	2.4	-	-
Plant matter	-	31.2	25.5	27.3	14.9	41.0	19.3
Unidentified matter	-	17.2	5.9	19.5	0.6	-	5.6

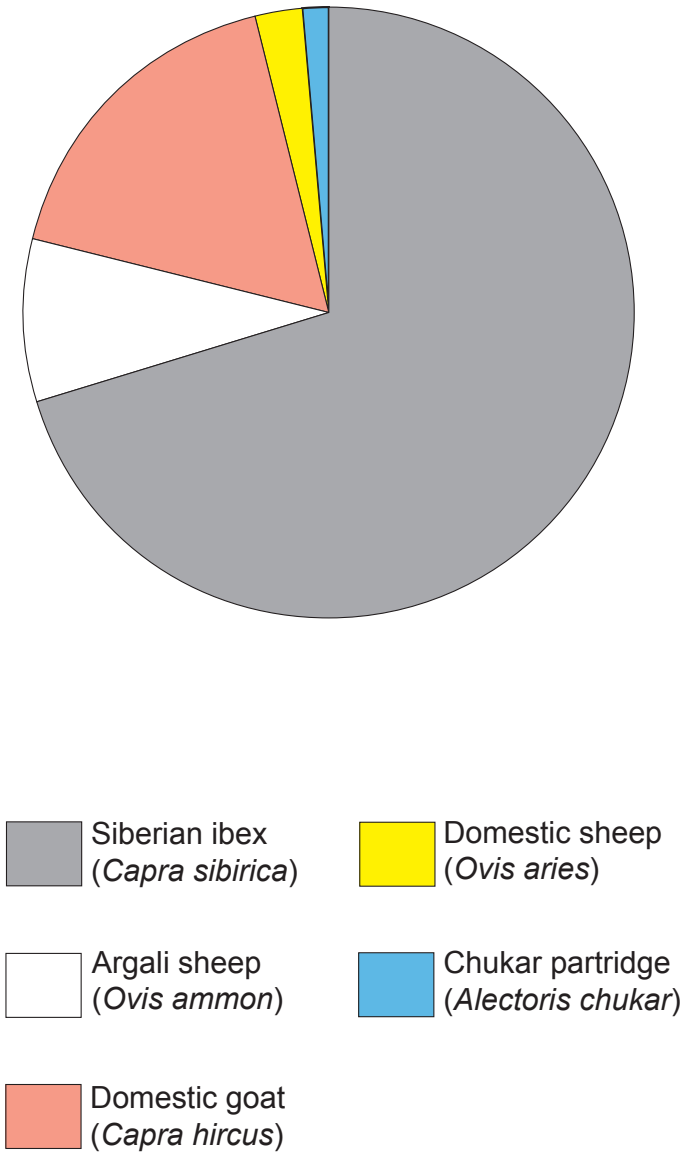
Table 4 Sequences of the primer pairs used in the study. The length of amplified fragments (excluding primers) with *Unci, 12SV5* and *Ovis* were 68 bp, ~100 bp and 82 bp, respectively.

Name	Primer sequence (5'-3')	Reference
UnciF	CTAAACCTAGATAGTTAGCT	Ficetola et al. 2010
UnciR	CTCCTCTAGAGGGGTG	Ficetola et al. 2010
12SV5F	TAGAACAGGCTCCTCTAG	Riaz et al. 2011
12SV5R	TTAGATACCCCACTATGC	Riaz et al. 2011
UnciB	CTATGCTTAGCCCTAAACCTAGATAGTTAG CTCAAACAAAACTAT-C3	This study
OvisF	AAACTATGGCTGAATTATCCGATA	This study
OvisR	TCCGATGTTTCATGTTTCTAGGAA	This study

Figure 1.

Figure 2

Prey frequency of snow leopard in Tost Mountains, South Gobi, Mongolia



6. Conclusions and perspectives

Combining the use of blocking oligonucleotides and universal primer pair for vertebrates with subsequent sequencing of PCR products with NGS has awesome potential to study carnivores diet. The blocking oligonucleotide specific to predator DNA restrains its amplification (e.g. 2.2 % of the total sequence count, in case of leopard cat diet) and at the same time allowing amplifying all vertebrate prey taxa present in a template, which otherwise could not be amplified with universal primers (see Fig. 3; article I). This technique has been presented by focusing primarily on a highly eclectic diet of the leopard cat (*Prionailurus bengalensis*). We were able to explore eighteen different prey taxa in 38 feces, belonging to different groups of mammals, birds, amphibians and fish, indicating that prey identification from a diverse range is possible with this method. Exploiting feces as a source of DNA, this system is particularly ideal to study the diet of endangered and elusive species, where non-invasive samples are the only source of information.

This system corresponds to a universal method and is highly adaptable to all carnivore species, by designing a blocking oligonucleotide homologous to predator DNA. We have applied this approach to other carnivore species i.e. the common leopard (*Panthera pardus*) and the snow leopard (*Panthera uncia*) to assess their diet (articles II and III). The availability of next generation sequencing together with the use of tagging systems that allow sorting out the sequences according to their sample of origin requires automated sequence analysis through the use of new bioinformatic tools. Thus, such approach is ideal for large-scale diet studies combining hundreds of samples together in a single sequencing run.

This system is robust and provides more precision for the identification of prey items consumed, as compared with the traditional methods of prey identification. We compared this approach with hairs microscopy using wolf feces (Shehzad *et al.* in prep; not presented in this manuscript). Our technique identified prey more accurately and precisely than microscopic examination that leads to ambiguities especially while identifying closely related species. When combined with the individual identification based on microsatellites, this system is able to assess individual diet as well as

differences related to sex, without monitoring them in the field. This methodology does not require any *a priori* information about the prey and is even cost effective.

An ongoing debate on NGS diet studies is their ability to quantify food items consumed through sequence counts obtained from PCR products characterization. This issue has been addressed in several recent studies (e.g. Deagle *et al.* 2009, 2010; Valentini *et al.* 2009a; Soininen *et al.* 2009). The sequence count cannot be interpreted as quantitative for several reasons. The biased amplification of some species has already been observed, when PCR with universal primers was carried out on a known mixture (Polz & Cavanaugh 1998). In dietary studies strong biases occur when mismatches exist between the primers and the sequence of certain prey, resulting in the preferential amplification of fragments bearing no mismatch on the priming sites. Biases may also occur due to differences in DNA fragments, especially when using environmental samples. In a recent study, Rayé *et al*. (2011) have reported that the short *trn*L size of *Helianthemum mummularium* compared to other plant species could be a reason of having more amplified sequences for this species. The two highly conserved regions targeted by the universal primers *12SV5F* and *12SV5R* make the approach less susceptible to PCR biases (article I). Deagle *et al*. (2010) suggested that differences in the tissue mitochondrial DNA density can also bias the sequence count. In the present study, we avoided interpreting our results in a quantitative way according to sequence counts, and we only recorded the presence/absence of the different prey in the different feces.

We address here certain limitations associated with DNA based diet studies in general and our proposed approach in particular with the probable solutions. Using feces or stomach contents to identify prey through DNA based methods provide a picture of the diet at a given time. However, in some cases long-term predation data are required for example for certain cryptic and endangered predators the long-term prey information is important for the conservation point of view. The diet inferred from isotopic signatures using tissue samples provides relatively long term information, but are unable to differentiate the prey taxa to species level. Another potential limitation of DNA based diet analysis is the impossibility to differentiate scavenging on dead animals or predation on alive prey. Scavenging may lead to false over estimation of the predation resulting in establishing wrong trophic interactions. Another point is that DNA based approaches are unable to determine the age of a prey, however few attempts have been made to study

the age of wild animals (see review by Dunshea *et al*. 2011). By exploiting the degradation of telomeres during the life to infer the age of an individual is not a reliable method. However, such information can only be inferred by monitoring of the animal in the field or by examining prey remains at the point of hunt.

Our method is unable to detect the prey in case of cannibalism when the predator has eaten its own young, as prey DNA cannot be distinguished from the predator DNA. Another limitation of the approach is that the *12SV5F* and *12SV5R* primers only amplify vertebrate DNA. Many carnivores have a more diverse diet including invertebrates and plants. For example, the Eurasian badger (*Meles meles*) exploits a wide range of food items with earthworms, insects, and grubs. It also eats small mammals, amphibians, reptiles and birds as well as roots and fruits (Revilla & Palomares 2002). For such kind of diet, we suggest to complement the universal primers for vertebrates with several additional systems, such as primer targeting plant taxa (e.g. Taberlet *et al.* 2007; Valentini *et al.* 2009a) or earthworms (Bienert *et al*. in press). Another potential difficulty concerns the identification of prey taxa especially, which are endemic to the study area and whose sequence is not present in public reference databases (see discussion of article I). We recommend building a local reference database for all the prey species present in the study area, for accurate assessment of diet.

The errors may occur during amplification process (i.e. PCR errors) or at the time of sequencing (i.e. sequencing errors). There may be two types of errors inherent due to PCR, chimeric products and point mutations (Acinas *et al.* 2005). The chances of PCR errors are even high when using environmental samples having DNA of mixed taxa, because this allows the formation of chimeric fragments during the PCR. Such errors result from the existence of incompletely amplified fragments or short fragments that act as primers in the next PCR cycles. We observed many of such chimeras in our sequencing files, but they were relatively few compared with the exact sequences having high copy numbers. Furthermore, they had low matching frequencies with sequences from reference datbases. The PCR induced point mutations occur due to misincorporation of bases. Such errors are attributed to *Taq* polymerase enzyme (Cline *et al.* 1996). If such error occurs during the early cycles of PCR, this will exponentially replicate these errors. As a result, we may have some closely related amplicons when compared with the reference sequences. Errors may also occur during the sequencing process due to misincorporation of bases. For example, the Illumina sequencing is

susceptible to substitutions rather than to insertions or deletions while the 454 sequencing is prone to sequencing errors like insertions or deletions and also having technical problem in sequencing of homopolymers. Such errors may make the species identification difficult. The error rates observed in 454 platform are more (~ 1 %) compared with Illumina platform (~0.1%) (Glenn 2011).

There is now a growing trend for molecular ecologists to study the diet. Despite some technical constraints, which still need to be addressed, like selection of metabarcodes and quantification of different prey in a sample, the DNA-based diet studies have a fabulous potential by allowing obtaining an accurate information less biased than that obtained with other techniques. Recent advances have successfully identified prey items consumed as in case of herbivores (Valentini *et al.* 2009), carnivores (present study) and omnivores (Alfstrom *et al.* in prep). Alternatively, where the use of PCR based metabarcodes is still problemantic, we may suggest to apply a capture probe approach, using oligonucleotides targeting conserved regions. These capture probes allow retrieving target fragments that will be sequenced using NGS for identifying the different taxa present in an environmental sample. The probability of finding a conserved probe is higher compared with that of finding PCR primer pair. In addition, many probes, targeting different taxonomic groups can be multiplexed in a single experiment to reveal all taxa (see Taberlet *et al*. submitted).

The diet analysis based upon the presence/absence of a prey is helpful to estimate diet composition. As discussed earlier in this chapter, it is difficult to obtain quantitative information, i.e. the proportion of the various prey items consumed, from such analyses. At the same time, DNA-based studies may fail to infer whether a predator has consumed one or several individuals of the same species. Without the ability to quantify the accurate diet estimates, the DNA-based diet studies would present some missing information. In a feeding trial experiment Deagle *et al.* (2010) have attempted to quantify the consumed prey species, but somehow, the data did not accurately reflect what predators have consumed. Trying to answer the same question, early experimentation is also in progress in our laboratory, by multiplexing various barcodes in a single PCR to study omnivores diet. In the future, we also suggest trying another approach to find the different proportions of various prey taxa present in an environmental sample using direct sequencing of environmental DNA extracts. The NGS platforms have ability to produce millions of sequences belonging to different taxa.

Even if a small part of these sequences would be identifiable by screening database, enough information would be available to describe the diversity of the sample. Furthermore, such a shotgun approach would have the potential advantage to identify the relative proportions among the various types of DNA originating from the different taxonomic groups (Taberlet *et al*. submitted).

The results presented in this manuscript reflect the diet as a result of an extensive sampling during one season. However, for studies addressing the conservation of endangered predators and prey species, we recommend long-term diet assessment. The concrete ecological conclusions can only be drawn with such long-term studies. Such studies are also be needed to predict accurate relationships between predators and prey, shift in prey species, etc.

References

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF (2005) PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology,* **71**, 8966-8969.

Agusti N, Shayler SP, Harwood JD, Vaughan IP, Sunderland KD, Symondson WOC (2003) Collembola as alternative prey sustaining spiders in arable ecosystems: prey detection within predators using molecular markers. *Molecular Ecology*, **12**, 3467-3475.

Ale SB, Brown JS (2009) Prey behavior leads to predator: a case study of the himalayan tahr and the snow leopard in Sagarmatha (Mt. Everest) National Park, Nepal. *Israel Journal of Ecology & Evolution,* **55**, 315-327.

Allison NL, Destefano S (2006) Equipment and techniques for nocturnal wildlife studies. *Wildlife Society Bulletin,* **34**, 1036-1044.

Ambrose SH, DeNiro MJ, (1986) The isotopic ecology of East African mammals. *Oecologia*, **69**, 395–406.

Anderson CR, Lindzey FG (2003) Estimating cougar predation rates from GPS location clusters. *Journal of Wildlife Management,* **67**, 307-316.

Ansorge WJ (2009) Next-generation DNA sequencing techniques. *New Biotechnology,* **25**, 195-203.

Anwar M, Jackson R, Nadeem M, Janečka J, Hussain S, Beg M, Muhammad G, Qayyum M (2011) Food habits of the snow leopard *Panthera uncia* (Schreber, 1775) in Baltistan, Northern Pakistan. *European Journal of Wildlife Research*, **57**, 1077-1083.

Arnot DE, Roper C, Bayoumi RAL (1993) Digital Codes from Hypervariable Tandemly Repeated DNA-Sequences in the Plasmodium-Falciparum Circumsporozoite Gene Can Genetically Barcode Isolates. *Molecular and Biochemical Parasitology,* **61**, 15-24.

Bagchi S, Mishra C (2006) Living with large carnivores: predation on livestock by the

snow leopard (Panthera uncia). *Journal of Zoology,* **268**, 217-224.

Balestrieri A, Remonti L, Prigioni C (2011) Assessing carnivore diet by faecal samples and stomach contents: a case study with Alpine red foxes. *Central European Journal of Biology,* **6**, 283-292.

Bauer H, Kari S (2001) Assessment of the people-predator conflict through thematic PRA in the surroundings of Waza National Park, Cameroon. *PLA Notes,* **41**, 9-13.

Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *Bmc Microbiology*, **10**, 9.

Bentley DR, Balasubramanian S, Swerdlow HP, *et al*. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**, 53-59.

Berth M, Delanghe J (2004) Protein precipitation as a possible important pitfall in the clinical chemistry analysis of blood samples containing monoclonal immunoglobulins: 2 case reports and a review of the literature. *Acta Clinica Belgica,* **59**, 263-273.

Bianchi RD, Rosa AF, Gatti A, Mendes SL (2011) Diet of margay, Leopardus wiedii, and jaguarundi, Puma yagouaroundi, (Carnivora: Felidae) in Atlantic Rainforest, Brazil. *Zoologia*, **28**, 127-132.

Bienert F, Sebastien de Danieli S, Miquel C, Eric Coissac E, Poillot C, Jean-Jacques Brun JJ, Taberlet P (2012) Tracking earthworm communities from soil DNA. *Molecular Ecology*, in press, doi: 10.1111/j.1365-294X.2011.05407.x

Blankenship LE, Yayanos AA (2005) Universal primers and PCR of gut contents to study marine invertebrate diets. *Molecular Ecology,* **14**, 891-899.

Bohmann K, Monadjem A, Noer CL, Rasmussen M, Zeale MRK, Clare E, Jones G, Willerslev E, Gilbert MTP (2011) Molecular Diet Analysis of Two African Free-Tailed Bats (Molossidae) Using High Throughput Sequencing. *PLoS ONE*, **6**, 11.

BowenWD, (2000) Reconstruction of pinniped diets: accounting for a complete digestion of otoliths and cephalopod beaks. *Canadian Journal of Fisheries and Aquatic Science* **57**, 898-905.

Braasch DA, Corey DR (2001) Locked nucleic acid (LNA): fine-tuning the recognition

of DNA and RNA. *Chemistry & Biology*, **8**, 1-7.

Braley M, Goldsworthy SD, Page B, Steer M, Austin JJ (2010) Assessing morphological and DNA-based diet analysis techniques in a generalist predator, the arrow squid Nototodarus gouldi. *Molecular Ecology Resources,* **10**, 466-474.

Britton JR, Pegg J, Shepherd JS, Toms S (2006) Revealing the prey items of the otter Lutra lutra in South West England using stomach contents analysis. *Folia Zoologica,* **55**, 167-174.

Chow S, Suzuki S, Matsunaga T, Lavery S, Jeffs A, Takeyama H (2011) Investigation on Natural Diets of Larval Marine Animals Using Peptide Nucleic Acid-Directed Polymerase Chain Reaction Clamping. *Marine Biotechnology*, **13**, 305-313.

Chun JY, Kim KJ, Hwang IT, Kim YJ, Lee DH, Lee IK, Kim JK (2007) Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. *Nucleic Acids Research*, **35**.

Chundawat RS, Rawat GS (1990) Food habits of the snow leopard in Ladakh. International Snow Leopard Trust, Seattle, USA.

Cline J, Braman JC, Hogrefe HH (1996) PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Research,* **24**, 3546-3551.

Corse E, Costedoat C, Chappaz R, Pech N, Martin JF, Gilles A (2010) A PCR-based method for diet analysis in freshwater organisms using 18S rDNA barcoding on faeces. *Molecular Ecology Resources*, **10**, 96-108.

Crooks KR, Soulé ME (1999) Mesopredator release and avifaunal extinctions in a fragmented system. *Nature,* **400**, 563-566.

Damm S, Schierwater B, Hadrys H (2010) An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology. *Molecular Ecology,* **19**, 3881-3893.

Dar NI, Minhas RA, Zaman Q, Linkie M (2009) Predicting the patterns, perceptions and causes of human-carnivore conflict in and around Machiara National Park, Pakistan. *Biological Conservation,* **142**, 2076-2082.

Darimont CT, Reimchen TE (2002) Intra-hair stable isotope analysis implies seasonal shift to salmon in gray wolf diet. *Canadian Journal of Zoology-Revue Canadienne*

De Zoologie, **80**, 1638-1642.

Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. *Biological Invasions,* **9**, 751-765.

Deagle BE (2006) *DNA-based Methods for Studying the Diet of Marine Predators*. PhD thesis, University of Tasmania.

Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics,* **11**, 2039-2048.

Deagle BE, Eveson JP, Jarman SN (2006) Quantification of damage in DNA recovered from highly degraded samples--a case study on DNA in faeces. *Front Zool,* **3**, 11.

Deagle BE, Gales NJ, Evans K, Jarman SN, Robinson S, Trebilco R, Hindell MA (2007a) Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (Eudyptes chrysolophus). *PLoS ONE*, **2**, e831.

Deagle BE, Jarman SN, Pemberton D, Gales NJ (2005a) Genetic screening for prey in the gut contents from a giant squid (*Architeuthis sp*.). *Journal of Heredity,* **96**, 417-423.

Deagle BE, Kirkwood R, Jarman SN (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology,* **18**, 2022-2038.

Deagle BE, Tollit DJ (2007b) Quantitative analysis of prey DNA in pinniped faeces: potential to estimate diet composition? *Conservation Genetics,* **8**, 743-747.

Deagle BE, Tollit DJ, Jarman SN, Hindell MA, Trites AW, Gales NJ (2005b) Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. *Molecular Ecology*, **14**, 1831-1842.

Demers DB, Curry ET, Egholm M, Sozer AC (1995) Enhanced PCR Amplification of VNTR locus D1S80 using Peptide Nucleic-Acid (PNA). *Nucleic Acids Research,* **23**, 3050-3055.

DeNiro MJ, Epstein S, (1978) Influence of diet on the distribution of carbon isotopes in animals. Geochim. Cosmochim. *Acta*, **42**, 495–506.

DeNiro MJ, Epstein S, (1981) Influence of diet on the distribution of nitrogen isotopes in animals. Geochim. Cosmochim. *Acta*, **45**, 341–351.

Dixon R, Coates D (2009) Near infrared spectroscopy of faeces to evaluate the nutrition and physiology of herbivores. *Journal of near Infrared Spectroscopy,* **17**, 1-31.

Dominguez PL, Kolodney MS (2005) Wild-type blocking polymerase chain reaction for detection of single nucleotide minority mutations from clinical specimens. *Oncogene,* **24**, 6830-6834.

Dunshea G (2009) DNA-Based Diet Analysis for Any Predator. *PLoS ONE,* **4**, e5252.

Dunshea G, Barros NB, Wells RS, Gales NJ, Hindell MA, Jarman SN (2008) Pseudogenes and DNA-based diet analyses: a cautionary tale from a relatively well sampled predator-prey system, 239-248.

Dunshea G, Duffield D, Gales N, Hindell M, Wells RS, Jarman SN (2011) Telomeres as age markers in vertebrate molecular ecology. *Molecular Ecology Resources*, **11**, 225-235.

Emrich CA, Tian H, Medintz IL, Mathies RA (2002) Microfabricated 384-Lane Capillary Array Electrophoresis Bioanalyzer for Ultrahigh-Throughput Genetic Analysis. *Analytical Chemistry,* **74**, 5076-5083.

Estes JA, Tinker MT, Williams TM, Doak DF (1998) Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science,* **282**, 473-476.

Feller RJ, Zagursky G, Day EA (1985) Deep-sea food web analysis using cross-reacting antisera. *Deep Sea Research Part A. Oceanographic Research Papers,* **32**, 485-497.

Feranec RS, MacFadden KBJ, (2000) Evolution of the grazing niche in Pleistocene mammals from Florida: evidence from stable isotopes. *Palaeogeography Palaeoclimatology Palaeoecology*, **162**, 155-169.

Fernandes CA, Ginja C, Pereira I, Tenreiro R, Bruford MW, Santos-Reis M (2008) Species-specific mitochondrial DNA markers for identification of non-invasive samples from sympatric carnivores in the Iberian Peninsula. *Conservation Genetics*, **9**, 681-690.

Fernandez GJ, Corley JC, Capurro AF (1997) Identification of cougar and jaguar feces through bile acid chromatography. *Journal of Wildlife Management,* **61**, 506-510.

Ficetola G, Coissac E, Zundel S, Riaz T, Shehzad W, Bessiere J, Taberlet P, Pompanon

F (2010) An In silico approach for the evaluation of DNA barcodes. *Bmc Genomics*, **11**, 434.

Foley WJ, McIlwee A, Lawler I, Aragones L, Wollnough AP, Berding N (1998) Ecological applications of near infrared reflectance spectroscopy - a tool for rapid, cost-effective prediction of the composition of plant and animal tissues ans aspects of animal performance. *Oecologia*, **116**, 293-305.

Fournier V, Hagler J, Daane K, de Leon J, Groves R (2008) Identifying the predator complex of Homalodisca vitripennis (Hemiptera : Cicadellidae): a comparative study of the efficacy of an ELISA and PCR gut content assay. *Oecologia,* **157**, 629-640.

Garcia NG, Feranec RS, Arsuaga JL, de Castro JMB, Carbonell E (2009) Isotopic analysis of the ecology of herbivores and carnivores from the Middle Pleistocene deposits of the Sierra De Atapuerca, northern Spain. *Journal of Archaeological Science,* **36**, 1142-1151.

Garros C, Ngugi N, Githeko AE, Tuno N, Yan G (2008) Gut content identification of larvae of the Anopheles gambiae complex in western Kenya using a barcoding approach. *Molecular Ecology Resources,* **8**, 512-518.

Gigli E, Rasmussen M, Civit S *et al.* (2009) An improved PCR method for endogenous DNA retrieval in contaminated Neandertal samples based on the use of blocking primers. *Journal of Archaeological Science,* **36**, 2676-2679.

Glenn TC (2011) Field guide to next-generation DNA sequencers. *Molecular Ecology Resources,* **11**, 759-769.

Gordon IJ (1994) Animal based measurement technique for grazing ecology research: a review. *Options Mediterraneennes*, 13-28.

Green SJ, Minz D (2005) Suicide polymerase endonuclease restriction, a novel technique for enhancing PCR amplification of minor DNA templates. *Applied and Environmental Microbiology,* **71**, 4721-4727.

Greenstone MH (1999) Spider Predation: How and why we study it. *The Journal of Arachnology,* **27**, 333-342.

Grönberg E (2011) *Movement patterns of snow leopard (Panthera uncia) around kills based on GPS location clusters*, Swedish University of Agricultural Sciences.

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. *PLoS ONE,* **6**, 7.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics,* **23**, 167-172.

Hall-Aspland S, Rogers T (2007) Identification of hairs found in leopard seal (Hydrurga leptonyx) scats. *Polar Biology,* **30**, 581-585.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identification through DNA barcodes. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **270**, 313-321.

Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. *Systematic Biology,* **54**, 852-859.

Hebert PDN, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **270**, S96-S99.

Hert DG, Fredlake CP, Barron AE (2008) Advantages and limitations of next-generation sequencing technologies: A comparison of electrophoresis and non-electrophoresis methods. *Electrophoresis,* **29**, 4618-4626.

Hobson KA, Clark RG (1992) Assessing avian diets using stable isotopes .2. factors influencing diet-tissue fractionation. *Condor,* **94**, 189-197.

Hollingsworth PM (2007) DNA barcoding: potential users. *Genomics, Society and Policy,* **3**, 44-47.

Horii T, Ohtsuka H, Osaki M, Ohkuni H (2009) Use of a dual priming oligonucleotide system to detect multiple sexually transmitted pathogens in clinical specimens. *Letters in Applied Microbiology,* **49**, 46-52.

Hoss M, Kohn M, Paabo S, Knauer F, Schroder W (1992) Excrement analysis by PCR. *Nature*, **359**, 199-199.

Hoyt M, Fleeger JW, Siebeling R, Feller RJ (2000) Serological estimation of prey-protein gut-residence time and quantification of meal size for grass shrimp

consuming meiofaunal copepods. *Journal of Experimental Marine Biology and Ecology*, **248**, 105-119.

Hu Y, Le Leu RK, Young GP (2009) Detection of K-ras mutations in azoxymethane-induced aberrant crypt foci in mice using LNA-mediated real-time PCR clamping and mutant-specific probes. *Mutation Research-Genetic Toxicology and Environmental Mutagenesis,* **677**, 27-32.

Hyrup B, Nielsen PE (1996) Peptide nucleic acids (PNA): Synthesis, properties and potential applications. *Bioorganic & Medicinal Chemistry,* **4**, 5-23.

Hyslop EJ (1980) Stomach contents analysis-a review of methods and their application. *Journal of Fish Biology,* **17**, 411-429.

Inskip C, Zimmermann A (2009) Review Human-felid conflict: a review of patterns and priorities worldwide. *Oryx,* **43**, 18-34.

Iverson SJ, Field C, Bowen WD, Blanchard W (2004) Quantitative fatty acid signature analysis: A new method of estimating predator diets. *Ecological Monographs,* **74**, 211-235.

Jackson RM (1996) *Home Range, Movements and Habitat Use of Snow Leopard (Uncia Uncia) In Nepal*. Ph.D. thesis, University of London.

Jarman SN, Deagle BE, Gales NJ (2004) Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology,* **13**, 1313-1322.

Jarman SN, Gales NJ, Tierney M, Gill PC, Elliott NG (2002) A DNA-based method for identification of krill species and its application to analysing the diet of marine vertebrate predators. *Molecular Ecology,* **11**, 2679-2690.

Jarman SN, Redd KS, Gales NJ (2006) Group-specific primers for amplifying DNA sequences that identify Amphipoda, Cephalopoda, Echinodermata, Gastropoda, Isopoda, Ostracoda and Thoracica. *Molecular Ecology Notes,* **6**, 268-271.

Jinbo U, Kato T, Ito M (2011) Current progress in DNA barcoding and future implications for entomology. *Entomological Science,* **14**, 107-124.

Juen A, Traugott M (2007) Revealing species-specific trophic links in soil food webs: molecular identification of scarab predators. *Molecular Ecology,* **16**, 1545-1557.

Kageyama T, Sato Y, Nishizawa S, Teramae N (2008) Competitive binding of small ligands to nucleobases in AP site-containing DNA duplexes. *Nucleic Acids Symposium Series*. **52**, 119-120.

Kaneko H, Lawler IR (2006) Can near infrared spectroscopy be used to improve assessment of marine mammal diets via fecal analysis? *Marine Mammal Science,* **22**, 261-275.

Knopff KH, Knopff AA, Warren MB, Boyce MS (2009) Evaluating Global Positioning System Telemetry Techniques for Estimating Cougar Predation Parameters. *Journal of Wildlife Management,* **73**, 586-597.

Kim BJ, Lee NS, Lee SD (2011) Feeding diets of the Korean water deer (Hydropotes inermis argyropus) based on a 202 bp rbcL sequence analysis. *Conservation Genetics,* **12**, 851-856.

King RA, Read DS, Traugott M, Symondson WOC (2008) Molecular analysis of predation: a review of best practice for DNA-based approaches. *Molecular Ecology,* **17**, 947-963.

King RA, Vaughan IP, Bell JR, Bohan DA, Symondson WOC (2010) Prey choice by carabid beetles feeding on an earthworm community analysed using species- and lineage-specific PCR primers. *Molecular Ecology,* **19**, 1721-1732.

Kircher M, Kelso J (2010) High-throughput DNA sequencing – concepts and limitations. *BioEssays,* **32**, 524-536.

Knudsen H, Nielsen PE (1997) Application of peptide nucleic acid in cancer therapy. *Anti-Cancer Drugs,* **8**, 113-118.

Kohn MH, Wayne RK (1997) Facts from feces revisited. *Trends in Ecology & Evolution,* **12**, 223-227.

Kowalczyk R, Taberlet P, Coissac E, Valentini A, Miquel C, Kaminski T, Wojcik JM (2011) Influence of management practices on large herbivore diet-Case of European bison in Bialowieza Primeval Forest (Poland). *Forest Ecology and Management*, **261**, 821-828.

Kuusk AK, Agusti N (2008) Group-specific primers for DNA-based detection of springtails (Hexapoda : Collembola) within predator gut contents. *Molecular Ecology Resources*, **8**, 678-681.

Kwak JY, Kim EK, Kim JK, Han JH, Hong SW, Park TS, Choi JR (2010) Dual Priming Oligonucleotide-Based Multiplex PCR analysis ror detection of Braf (V600E) mutation in Fnab samples of thyroid nodules in Braf (V600E) Mutation-prevalent area. *Head and Neck-Journal for the Sciences and Specialties of the Head and Neck*, **32**, 490-498.

Lecomte N, Ahlstrom O, Ehrich D, Fuglei E, Ims RA, Yoccoz NG (2011) Intrapopulation Variability Shaping Isotope Discrimination and Turnover: Experimental Evidence in Arctic Foxes. *PLoS ONE*, **6**, 10.

Lhagvasuren B, Munkhtsog B (2000) The yak population in Mongolia and its relation with snow leopards as a prey species. In: *Proceedings of the Third International Congress on Yak production in central Asian highlands* (eds Jianlin H, Richard C, Hanotte O, McVeigh C, Rege JEO), pp. 69-75. International Livestock Research Institute, Lhasa, China.

Li XM, Ng MTT, Wang YF, Liu XQ, Li TH (2007) Dumbbell-shaped circular oligonucleotides as inhibitors of human topoisomerase I. *Bioorganic & Medicinal Chemistry Letters,* **17**, 4967-4971.

Li MJ, Sato Y, Nishizawa S, Seino T, Nakamura K, Teramae N (2009) 2-Aminopurine-Modified Abasic-Site-Containing Duplex DNA for Highly Selective Detection of Theophylline. *Journal of the American Chemical Society*, **131**, 2448-2449.

Liles MR, Manske BF, Bintrim SB, Handelsman J, Goodman RM (2003) A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Applied and Environmental Microbiology,* **69**, 2684-2691.

Macadangdang B, Zhang N, Lund PE, Marple AH, Okabe M, Gottesman MM, Appella DH, Kimchi-Sarfaty C (2011) Inhibition of Multidrug Resistance by SV40 Pseudovirion Delivery of an Antigene Peptide Nucleic Acid (PNA) in Cultured Cells. *PLoS ONE*, **6**, 6.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen ZT, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis

GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu PG, Begley RF, Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**, 376-380.

Marshall HD, Hart KA, Yaskowiak ES, Stenson GB, McKinnon D, Perry EA (2010) Molecular identification of prey in the stomach contents of Harp Seals (Pagophilus groenlandicus) using species-specific oligonucleotides. *Molecular Ecology Resources*, **10**, 181-189.

McCarthy KP, Fuller TK, Ming M, McCarthy TM, Waits L, Jumabaev K (2008) Assessing Estimators of Snow Leopard Abundance. *Journal of Wildlife Management*, **72**, 1826-1833.

Merfield CN, Wratten SD, Navntoft S (2004) Video analysis of predation by polyphagous invertebrate predators in the laboratory and field. *Biological Control,* **29**, 5-13.

Metzker ML (2010) Applications of Next-Generation sequencing, Sequencing technologies - the next generation. *Nature Reviews Genetics,* **11**, 31-46.

Metzler BU, Vahjen W, Baumgartel T, Rodehutscord M, Mosenthin R (2009) Changes in bacterial populations in the ileum of pigs fed low-phosphorus diets supplemented with different sources of fermentable carbohydrates. *Animal Feed Science and Technology*, **148**, 70-91.

Mills MGL (1992) A comparison of methods used to study food habits of large African carnivores. In: *Wildlife 2001: Populations* (eds. D. R. McCullough, H. Barrett). Elsevier Science Publishers, London.

Mills LS, Soule ME, Doak DF (1993) The keystone-species concept in ecology and conservation. *Bioscience,* **43**, 219-224.

Mishra C (1997) Livestock depredation by large carnivores in the Indian trans-Himalaya: conflict perceptions and conservation prospects. *Environmental Conservation,* **24**, 338-343.

Moyo T (2005) *The identification of mammalian species through the classification of hair patterns using image pattern recognition*. Bachelor of Science (Honors), Rhodes University.

Mukherjee S, Goyal SP, Chellam R (1994) Refined techniques for the analysis of asiatic lion *panthera leo persica* scats. *Acta Theriologica,* **39**, 425-430.

Najafi-Shoushtari SH, Kristo F, Li YX, Shioda T, Cohen DE, Gerszten RE, Naar AM (2010) MicroRNA-33 and the SREBP Host Genes Cooperate to Control Cholesterol Homeostasis. *Science*, **328**, 1566-1569.

Namgail T, Fox JL, Bhatnagar YV (2007) Carnivore-caused livestock mortality in Trans-Himalaya. *Environmental Management,* **39**, 490-496.

Naranjo SE, Hagler JR (2001) Toward the quantification of predation with predator gut immunoassays: A new approach integrating functional response behavior. *Biological Control*, **20**, 175-189.

Nijboer J, Dierenfeld ES (1996) Comparison of diets fed to south-east Asian colobines in North America and European zoos, with emphasis on temperate browse composition. *Zoo Biology*, **15**, 499-507.

Nielsen PE, Egholm M (1999) An Introduction to Peptide Nucleic Acid. *Current issues in Molecular Biology,* **1**, 89-104.

Oehm J, Juen A, Nagiller K, Neuhauser S, Traugott M (2011) Molecular scatology: how to improve prey DNA detection success in avian faeces? *Molecular Ecology Resources,* **11**, 620-628.

Ogra MV (2008) Human-wildlife conflict and gender in protected area borderlands: A case study of costs, perceptions, and vulnerabilities from Uttarakhand (Uttaranchal), India. *Geoforum,* **39**, 1408-1422.

Oli MK (1993) A key for the identification of the hair of mammals of a snow leopard (*Pnathera uncia*) habitat in Nepal *Journal of Zoology,* **231**, 71-93.

Ørum H (2000) PCR Clamping. *Current Issues in Molecular Biology,* **2**, 27-30.

Ørum H, Nielsen PE, Egholm M *et al.* (1993) Single-Base Pair Mutation analysis by PNA Directed PCR Clamping. *Nucleic Acids Research,* **21**, 5332-5336.

Osborne BG, Fearn T, Hindle PH (1993) Practical NIR spectroscopy with applications in food and beverage analysis. *Longman*, Singapore.

Park RS, Agnew RE, Gordon FJ, Steen RWJ (1998) The use of near infrared reflectance spectroscopy (NIRS) on undried samples of grass silage to predict chemical

composition and digestibility parameters. *Animal Feed Science and Technology,* **72**, 155-167.

Pegard A, Miquel C, Valentini A, Coissac E, Bouvier F, Francois D, Taberlet P, Engel E, Pompanon F (2009) Universal DNA-Based Methods for Assessing the Diet of Grazing Livestock and Wildlife from Feces. *Journal of Agricultural and Food Chemistry*, **57**, 5700-5706.

Pérez-Claros JA, Palmqvist P (2008) How many potential prey species account for the bulk of the diet of mammalian predators? Implications for stable isotope paleodietary analyses. *Journal of Zoology,* **275**, 9-17.

Piche J, Iverson SJ, Parrish FA, Dollar R (2010) Characterization of forage fish and invertebrates in the Northwestern Hawaiian Islands using fatty acid signatures: species and ecological groups. *Marine Ecology-Progress Series,* **418**, 1-15.

Pickering TR (2001) Carnivore voiding: A taphonomic process with the potential for the deposition of forensic evidence. *Journal of Forensic Sciences,* **46**, 406-411.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, **64**, 3724-3730.

Pompanon F, Deagle BE, Symondson WOC, Brown DS, Jarman SN, Taberlet P (2012) Who is eating what: diet assessment using Next Generation Sequencing. *Molecular ecology,* in press doi: 10.1111/j.1365-294X.2011.05403.x.

Power ME, Tilman D, Estes JA, Menge BA, Bond WJ, Mills LS, Daily G, Castilla JC, Lubchenco J, Paine RT (1996) Challenges in the quest for keystones. *Bioscience*, **46**, 609-620.

Prado P, Alcoverro T, Romero J (2010) Influence of nutrients in the feeding ecology of seagrass (*Posidonia oceanica* L.) consumers: a stable isotopes approach. *Marine Biology,* **157**, 715-724.

Putman RJ (1984) Facts from feces. *Mammal Review,* **14**, 79-97.

Quinn T, Jackman WR (1994) Influence of diet on detection of fecal bile-acids by thin-layer chromatography. *Journal of Wildlife Management,* **58**, 295-299.

Raye G, Miquel C, Coissac E, Redjadj C, Loison A, Taberlet P (2011) New insights on diet variability revealed by DNA barcoding and high-throughput pyrosequencing:

chamois diet in autumn as a case study. *Ecological Research*, **26**, 265-276.

Read DS, Sheppard SK, Bruford MW, Glen DM, Symondson WOC (2006) Molecular detection of predation by soil micro-arthropods on nematodes. *Molecular Ecology,* **15**, 1963-1972.

Revilla E, Palomares F (2002) Does local feeding specialization exist in Eurasian badgers? *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **80**, 83-93.

Riaz , Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research,* **39**, 11.

Robbins CT, Felicetti LA, Sponheimer M (2005) The effect of dietary protein quality on nitrogen isotope discrimination in mammals and birds. *Oecologia,* **144**, 534-540.

Roth JD, Hobson KA (2000) Stable carbon and nitrogen isotopic fractionation between diet and tissue of captive red fox: implications for dietary reconstruction. *Canadian Journal of Zoology-Revue Canadienne De Zoologie,* **78**, 848-852.

Rundel PW, Ehleringer JR, Nagy K (1989) Isotopes in ecological research. *Springer-Verlag*, Berlin.

Sand H, Zimmermann B, Wabakken P, Andren H, Pedersen HC (2005) Using GPS technology and GIS cluster analyses to estimate kill rates in wolf-ungulate ecosystems. *Wildlife Society Bulletin,* **33**, 914-925.

Sangay T, Vernes K (2008) Human-wildlife conflict in the Kingdom of Bhutan: Patterns of livestock predation by large mammalian carnivores. *Biological Conservation,* **141**, 1272-1282.

Sawata S, Kai E, Ikebukuro K, Iida T, Honda T, Karube I (1999) Application of peptide nucleic acid to the direct detection of deoxyribonucleic acid amplified by polymerase chain reaction. *Biosensors & Bioelectronics*, **14**, 397-404.

Schadt EE, Turner S, Kasarskis A (2010) A window into third-generation sequencing. *Human Molecular Genetics,* **19**, R227-R240.

Scheiffarth G (2001) The diet of bar-tailed godwits *Limosa lapponica* in the wadden sea: combining visual observations and faeces analyses. *Adrea,* **89**, 481-494.

Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nature*

Methods, **5**, 16-18.

Sergio F, Caro T, Brown D, Clucas B, Hunter J, Ketchum J, McHugh K, Hiraldo F (2008) Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. *Annual Review of Ecology Evolution and Systematics*, **39**, 1-19.

Shehzad W, McCarthy TM, Pompanon F, Purevjav L, Coissac E, Riaz T, Taberlet P. Prey preference of snow leopard (*Panthera uncia*) in South Gobi, Mongolia. *PLoS ONE,* in revision.

Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P (2012) Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology*, in press, doi: 10.1111/j.1365-294X.2011.05424.x

Sheikh KM, Molur S (2004) Status and Red List of Pakistan's Mammals. Based on the Conservation Assessment and Management Plan Workshop. IUCN Pakistan.

Shendure J, Ji HL (2008) Next-generation DNA sequencing. *Nature Biotechnology,* **26**, 1135-1145.

Sheppard SK, Harwood JD (2005) Advances in molecular ecology: tracking trophic links through predator–prey food-webs. *Functional Ecology*, **19**, 751-762.

Shibata K, Itoh M, Aizawa K, Nagaoka S, Sasaki N, Carninci P, Konno H, Akiyama J, Nishi K, Kitsunai T, Tashiro H, Itoh M, Sumi N, Ishii Y, Nakamura S, Hazama M, Nishine T, Harada A, Yamamoto R, Matsumoto H, Sakaguchi S, Ikegami T, Kashiwagi K, Fujiwake S, Inoue K, Togawa Y, Izawa M, Ohara E, Watahiki M, Yoneda Y, Ishikawa T, Ozawa K, Tanaka T, Matsuura S, Kawai J, Okazaki Y, Muramatsu M, Inoue Y, Kira A, Hayashizaki Y (2000) RIKEN Integrated Sequence Analysis (RISA) System—384-Format Sequencing Pipeline with 384 Multicapillary Sequencer. *Genome Research*, **10**, 1757-1771.

Shrestha R, Wegge P (2006) Determining the composition of herbivore diets in the trans-Himalayan rangelands: A comparison of field methods. *Rangeland Ecology & Management,* **59**, 512-518.

Soininen EM, Valentini A, Coissac E, Miquel C, Gielly L, Brochmann C, Brysting AK, Sonstebo JH, Ims RA, Yoccoz NG, Taberlet P (2009) Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput

pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in Zoology*, **6**, 9.

Stech M, Kolvoort E, Loonen M, Vrieling K, Kruijer JD (2011) Bryophyte DNA sequences from faeces of an arctic herbivore, barnacle goose (*Branta leucopsis*). *Molecular Ecology Resources,* **11**, 404-408.

Suzuki N, Hoshino K, Murakami K, Takeyama H, Chow S (2008) Molecular diet analysis of phyllosoma larvae of the Japanese spiny lobster Panulirus japonicus (Decapoda : Crustacea). *Marine Biotechnology,* **10**, 49-55.

Suzuki N, Murakami K, Takeyama H, Chow S (2006) Molecular attempt to identify prey organisms of lobster phyllosoma larvae. *Fisheries Science,* **72**, 342-349.

Symondson WOC (2002) Molecular identification of prey in predator diets. *Molecular Ecology,* **11**, 627-641.

Symondson WOC, Erickson ML, Liddell JE (1999) Development of a monoclonal antibody for the detection and quantification of predation on slugs within the Arion hortensis agg. (Mollusca : Pulmonata). *Biological Control*, **16**, 274-282.

Symondson WOC, Sunderland KD, Greenstone MH (2002) Can generalist predators be effective biocontrol agents? *Annual Review of Entomology*, **47**, 561-594.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (submitted) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, **35**, 8.

Taberlet P, Fumagalli L (1996) Owl pellets as a source of DNA for genetic studies of small mammals. *Molecular Ecology,* **5**, 301-305.

Taberet P, Prud'homme S, Campione E, Roy J, Miquel C, Shehzad W, Gielly L, Rioux D, Melo de Lima C, Pompanon F, Coissac E (2012) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, in press, Doi: 10.1111/j.1365-294X.2011.05317.x

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution*, **14**, 321-325.

Tieszen LL, Boutton TW, Tesdahl KG, Slade NA, (1983) Fractionation and turnover of stable carbon isotopes in animal tissues: implications for δ13C analysis of diet. *Oecologia*, **57**, 32–37.

Tollit DJ, Wong M, Winship AJ, Rosen DAS, Trites AW (2003) Quantifying errors associated with using prey skeletal structures from fecal samples to determine the diet of Steller's sea lion (*Eumetopias jubatus*). *Marine Mammal Science,* **19**, 724-744.

Tollit DJ, Schulze AD, Trites AW, Olesiuk PF, Crockford SJ, Gelatt TS, Ream RR, Miller KM (2009) Development and application of DNA techniques for validating and improving pinniped diet estimates. *Ecological Applications*, **19**, 889-905.

Tran H, Salgado P, Tillard E, Dardenne P, Nguyen XT, Lecomte P (2010) "Global" and "local" predictions of dairy diet nutritional quality using near infrared reflectance spectroscopy. *Journal of Dairy Science*, **93**, 4961-4975.

Troedsson C, Simonelli P, Nagele V, Nejstgaard JC, Frischer ME (2009) Quantification of copepod gut content by differential length amplification quantitative PCR (dla-qPCR). *Marine Biology,* **156**, 253-259.

Tsukahara T (1993) Lions eat chimpanzees - The 1st evidence of predation by lions on wild chimpanzees. *American Journal of Primatology,* **29**, 1-11.

Traugott M (2003) The prey spectrum of larval and adult Cantharis species in arable land: An electrophoretic approach. *Pedobiologia,* **47**, 161-169.

Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, Pompanon F, Gielly L, Cruaud C, Nascetti G, Wincker P, Swenson JE, Taberlet P (2009a) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trn*L approach. *Molecular Ecology Resources*, **9**, 51-60.

Valentini A, Pompanon F, Taberlet P (2009b) DNA barcoding for ecologists. *Trends in Ecology & Evolution,* **24**, 110-117.

Varley MJ, Copland MJW, Wratten SD, Bowie MH (1994) Parasites and predators. In: *Video Techniques in Animal Ecology and Behaviour* (ed. Wratten SD), pp. 33–63. Chapman & Hall, London.

Vestheim H, Edvardsen B, Kaartvedt S (2005) Assessing feeding of a carnivorous copepod using species-specific PCR. *Marine Biology,* **147**, 381-385.

Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples - a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology,* **5**, 11.

Waits LP, Paetkau D (2005) Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. *Journal of Wildlife Management,* **69**, 1419-1433.

Walrant A, Loreau M (1995) Comparison of iso-enzyme electrophoresis and gut content examination for determining the natural diets of the groundbeetle species Abax ater (Coleoptera: Carabidae). *Entomolgia Generalis*, **19**, 253–259.

Walter J, Tannock GW, Tilsala-Timisjarvi A, Rodtong S, Loach DM, Munro K, Alatossava T (2000) Detection and identification of gastrointestinal Lactobacillus species by using denaturing gradient gel electrophoresis and species-specific PCR primers. *Applied and Environmental Microbiology*, **66**, 297-303.

Wang YF, Ng MTT, Zhou TY, Li XM, Tan CH, Li TH (2008) C3-Spacer-containing circular oligonucleotides as inhibitors of human topoisomerase I. *Bioorganic & Medicinal Chemistry Letters*, **18**, 3597-3602.

Wilson RP (1984) An improved stomach pump for penguines and other seabirds. *Journal of Field Ornithology,* **55**, 109-112.

Woo HY, Park H, Kim BI, Jeon WK, Kim YJ (2008) Evaluation of dual priming oligonucleotide-based multiplex PCR for detection of HBV YMDD mutants. *Archives of Virology,* **153**, 2019-2025.

Yoo SJ, Kuak EY, Shin BM (2007) Detection of 12 respiratory viruses with two-set multiplex reverse transcriptase-PCR assay using a dual priming oligonucleotide system. *Korean Journal of Laboratory Medicine*, **27**, 420-427.

Zeale MRK, Butlin RK, Barker GLA, Lees DC, Jones G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources,* **11**, 236-244